(NASA-TM-X-58119) THE ERTS-1 INVESTIGATION (ER-600). VOLUME 3: ERTS-1 FOREST ANALYSIS Technical Report, Jul. 1972 - Jun. 1973 (NASA) 105 p HC $5.50 CSCL 02F

N76-11536

Unclas
G3/43 01568

JSC-08458

NASA TECHNICAL MEMORANDUM

NASA TM X-58119
June 1974

THE ERTS-1 INVESTIGATION (ER-600)

VOLUME III — ERTS-1 FOREST ANALYSIS

(REPORT FOR PERIOD JULY 1972 - JUNE 1973)

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

LYNDON B. JOHNSON SPACE CENTER

HOUSTON, TEXAS 77058

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM X-58119		

4. Title and Subtitle	5. Report Date
THE ERTS-1 INVESTIGATION (ER-600): VOLUME III — ERTS-1 FOREST ANALYSIS (REPORT FOR PERIOD JULY 1972 - JUNE 1973)	June 1974
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
R. Bryan Erb	JSC-08458
	10. Work Unit No.
9. Performing Organization Name and Address	641-14-07-50-72
Lyndon B. Johnson Space Center Houston, Texas 77058	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Memorandum
National Aeronautics and Space Administration Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

The JSC Director waived the use of the International System of Units (SI) for this Technical Memorandum because, in his judgment, the use of SI Units would impair the usefulness of the report or result in excessive cost.

16. Abstract

The Forest Analysis Team of the Lyndon B. Johnson Space Center Earth Observations Division conducted a year's investigation of the Earth Resources Technology Satellite (ERTS-1) multi-spectral data to determine the size of forest features that could be detected and to determine the suitability for making forest classification maps. The Sam Houston National Forest of Texas was used as the test site. Using conventional image interpretation and computer-aided techniques, the team was able to differentiate up to 14 classes of forest features to an accuracy ranging between 55 and 84 percent.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
· Vegetation · Timber · National Forests · Forest Classification	STAR Subject Category: ~~18~~ 43

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	104	

*For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151

PREFACE

This report is one of seven separate reports prepared by six discipline-oriented analysis teams of the Earth Observations Division at the NASA Lyndon B. Johnson Space Center, Houston, Texas.

The seven reports were prepared originally for Goddard Space Flight Center in compliance with requirements for the Earth Resources Technology Satellite (ERTS-1) Investigation (ER-600). The project was approved and funded by NASA Headquarters in July 1972.

This report (Volume III) was accomplished by the Forest Analysis Team. The following is a list of the team members.

N. J. Clinton, Lockheed Electronics Co., Inc.
G. R. Heath, Lockheed Electronics Co., Inc.
W. A. Kunkle, Lockheed Electronics Co., Inc.
R. A. Pinter, Lockheed Electronics Co., Inc.
J. E. Weaver, Lockheed Electronics Co., Inc.

The total investigation is documented in the following reports.

Volume	Title	NASA Number
	A COMPENDIUM OF ANALYSIS RESULTS OF THE UTILITY OF ERTS-1 DATA FOR LAND RESOURCES MANAGEMENT	SP-347 JSC-08455
I	ERTS-1 AGRICULTURAL ANALYSIS	TM X-58117 JSC-08456
II	ERTS-1 COASTAL/ESTUARINE ANALYSIS	TM X-58118 JSC-08457

Volume	Title	NASA Number
III	ERTS-1 FOREST ANALYSIS	TM X-58119 JSC-08458
IV	ERTS-1 RANGE ANALYSIS	TM X-58120 JSC-08459
V	ERTS-1 URBAN LAND USE ANALYSIS	TM X-58121 JSC-08460
VI	ERTS-1 SIGNATURE EXTENSION ANALYSIS	TM X-58122 JSC-08461
VII	ERTS-1 LAND-USE ANALYSIS OF THE HOUSTON AREA TEST SITE	TM X-58124 JSC-08463

CONTENTS

Section		Page
1.0	SUMMARY	1-1
1.1	GENERAL	1-1
1.2	OBJECTIVES	1-2
1.3	GROUND TRUTH	1-2
1.4	CLASSIFICATION FEATURES	1-3
1.5	ANALYSIS APPROACH	1-3
1.6	EVALUATION OF DETECTABLE AREA SIZE	1-3
1.7	CLASSIFICATION BY CONVENTIONAL METHODS	1-5
	1.7.1 Single-Band Conventional Classification	1-5
	1.7.2 Multiband Conventional Methods	1-6
1.8	CLASSIFICATION BY COMPUTER METHODS	1-7
	1.8.1 Clustering Classification Method	1-7
	1.8.2 Maximum-Likelihood Classification Method	1-7
1.9	SIGNATURE EXTENSION	1-8
1.10	OTHER SIGNIFICANT FINDINGS	1-8
1.11	ACCURACY AND ANALYSIS	1-9
1.12	DISCUSSION	1-10
1.13	CONCLUSIONS	1-11
1.14	RECOMMENDATIONS	1-12
2.0	INTRODUCTION	2-1
2.1	GENERAL	2-1
2.2	OBJECTIVES	2-1

Section		Page
3.0	STUDY AREA, FEATURES, AND CRITERIA	3-1
3.1	USER REQUIREMENTS.	3-1
3.2	STUDY AREA SELECTION	3-1
3.3	FEATURES	3-2
3.4	CRITERIA	3-7
4.0	DATA UTILIZATION	4-1
4.1	ERTS-1 DATA.	4-1
4.2	AIRCRAFT DATA.	4-1
4.3	GROUND SURVEY.	4-3
5.0	ANALYSIS APPROACH.	5-1
5.1	SENSOR DETECTION EVALUATION.	5-1
5.2	CONVENTIONAL IMAGERY INTERPRETATION PROCESSING	5-1
	5.2.1 Single-Band Classification.	5-2
	5.2.2 Multiband Classification.	5-2
5.3	CLASSIFICATION BY COMPUTER METHODS	5-3
	5.3.1 Clustering Method	5-4
	5.3.2 Maximum-Likelihood Method	5-5
5.4	SIGNATURE EXTENSION.	5-6
	5.4.1 Conventional Classification Extension	5-6
	5.4.2 Computer Classification	5-7
5.5	ACCURACY EVALUATION.	5-7
6.0	ANALYSIS RESULTS	6-1
6.1	EVALUATION OF DETECTABLE AREA RESULTS. . . .	6-1

Section		Page
6.2	RESULTS OF CLASSIFICATION BY CONVENTIONAL METHODS	6-6
	6.2.1 Single-Band Classification Results	6-6
	6.2.2 Multiband Classification Results. . .	6-6
6.3	RESULTS OF CLASSIFICATION BY COMPUTER METHODS	6-11
	6.3.1 Clustering Results.	6-11
	6.3.2 Maximum-Likelihood Classification Results.	6-21
6.4	SIGNATURE EXTENSION RESULTS.	6-27
	6.4.1 Conventional Classification Results	6-27
	6.4.2 Computer Classification Results . . .	6-28
	6.4.3 Information Extension Results	6-31
6.5	ACCURACY EVALUATION RESULTS.	6-34
	6.5.1 Accuracy Evaluation Results	6-34
7.0	ADDITIONAL EXPERIMENTAL WORK	7-1
8.0	DISCUSSION OF TECHNIQUE DEVELOPMENT.	8-1
9.0	CONCLUSIONS AND RECOMMENDATIONS.	9-1
9.1	CONCLUSIONS.	9-1
9.2	RECOMMENDATIONS.	9-2
10.0	REFERENCES	10-1
APPENDIX A - OTHER SIGNIFICANT FINDINGS		A-1
APPENDIX B - GLOSSARY		B-2

TABLES

Table		Page
1-I	NUMBER OF TARGETS DETECTED	1-5
1-II	ACCURACY	1-10
3-I	HIERARCHY OF FOREST FEATURES	3-2
6-I	MINIMUM SIZE OF FOREST FEATURES	6-2
6-II	TYPES OF DETECTION TARGETS	6-3
6-III	PERCENT OF TARGETS DETECTED	6-4
6-IV	NUMBER OF TARGETS DETECTED	6-5
6-V	SUMMARY OF ACCURACIES OF CLASSIFICATIONS USED	6-33
6-VI	ACCURACY OF INDIVIDUAL CLASSES	6-35
A-I	COMPUTATIONS OF ACCURACY BY THE AREA DIFFERENCE METHOD IN THE PRIMARY STUDY AREA	A-4
A-II	COMPUTATIONS OF ACCURACY BY THE AREA DIFFERENCE METHOD IN THE SECONDARY STUDY AREA	A-5
A-III	STANDARD DEVIATIONS OF THE CLUSTERS AND THE DISTANCES BETWEEN CLUSTER MEANS OF THE FIRST NINE CLUSTERS OF THE LARGE-AREA CLUSTERING CLASSIFICATION	A-6
A-IV	PARTIAL RESULTS OF DISTANCES BETWEEN CLUSTERS IN LARGE-AREA CLUSTERING CLASSIFICATION	A-7
A-V	STATISTICS SHOWING THE STANDARD DEVIATIONS OF THE CLUSTERS AND DISTANCES BETWEEN CLUSTER MEANS OF THE SMALL-AREA CLUSTERING CLASSIFICATION	A-8
A-VI	PARTIAL RESULTS OF DISTANCES BETWEEN CLUSTERS IN SMALL-AREA CLUSTERING CLASSIFICATION	A-9

Table | Page

A-VII LARSYS TRAINING FIELDS A-10

A-VIII STATISTICS OF LARSYS TEST FIELDS A-11

FIGURES

Figure		Page
2-1	Location map of ERTS-1 forestry study area	2-2
4-1	Ground-truth classification map of primary area	4-2
6-1	Single-band conventional classification map of primary area	6-7
6-2	Optical multiband classification map of primary area	6-8
6-3	MCFV classification map of primary area	6-10
6-4	JSC color composite classification map of primary area	6-12
6-5	Improved JSC color composite of the study area	6-13
6-6	Computer printout of cluster classification of primary area	6-15
6-7	Classification map of primary area derived from clustering	6-18
6-8	DAS recording of cluster classification map of primary area	6-20
6-9	DAS recording of a small-area cluster classification map of the primary area	6-22
6-10	LARSYS classification map of primary area	6-25
6-11	DAS recording of LARSYS classification map of primary area	6-26
6-12	Single-band conventional classification map of secondary area	6-29

Figure		Page
6-13	Ground-truth classification map of secondary area	6-30
6-14	Classification map of secondary area derived from clustering	6-32
A-1	JSC enhancement of light ground fire in the Sam Houston National Forest	A-2

ABBREVIATIONS AND ACRONYMS

Acc.	Accuracy
CIR	color infrared
cm	centimeter
DLMIN	Minimum distance between cluster means
EBR	Electron beam recorder
ERIPS	Earth Resources Interactive Processing System
ERTS	Earth Resources Technology Satellite
ft	feet
GSFC	Goddard Space Flight Center
hr	hour
in	inch
IR	infrared
JSC	Lyndon B. Johnson Space Center
LARS	Laboratory for Applications of Remote Sensing
LARSYS	Maximum-likelihood classification algorithm
MCFV	Multichannel film viewer
mm	millimeter
MSS	multispectral scanner
No.	number
R/W	rights of way
SHNF	Sam Houston National Forest
STDMAX	maximum standard deviation

THE ERT-1 INVESTIGATION (ER-600)

VOLUME III - ERTS-1 FOREST ANALYSIS

(REPORT FOR PERIOD JULY 1972 - JUNE 1973)

By R. Bryan Erb
Lyndon B. Johnson Space Center

1.0 SUMMARY

1.1 GENERAL

The Earth Resources Technology Satellite (ERTS-1) data were evaluated for their applications to forestry in two related studies; on a primary area where ground truth existed, and on a secondary area where none existed. A study was made to determine the size of forest features detectable under varying conditions. Computer-generated composite imagery from three spectral bands produced the best detectability for forest stands down to 4 hectares. Conventional imagery interpretation, as well as computer processing methods, were used in another study to make classification maps of timber species, condition, and land use. The results were compared with timber-stand maps obtained from aircraft imagery and checked in the field. The accuracy achieved ranged from 60 percent to 74 percent for four types of conventional and two types of computer classifications. Computerized analysis classified three major forest features with a mean of 82-percent accuracy. It is concluded that ERTS-1 data can probably be best used in forestry application where a few broad generalized classes are needed rather than where up to 14 classes should be identified with accuracies of at least 80 percent. This capability suggests that computer analysis can be used for differentiation of Levels 1 and 2 (described on page 1-3) for forest surveys with accuracies of better than 80 percent.

This permits new concepts of large area surveys to be developed.

The Sam Houston National Forest, Texas, were selected as the test site because of the knowledge of the area and its easy accessibility for field checks. The area comprised a primary study area surrounded by a secondary study area. The primary study area was the training area for methods and techniques development, and the secondary study area was used to verify these methods and techniques.

Only one set of ERTS-1 data was analyzed in this investigation, that of August 29, 1972.

1.2 OBJECTIVES

The objectives of the investigation were (1) to determine the size of the forest features which could be detected in the data under varying conditions, (2) to determine the suitability of the data for making forest classification maps, (3) to test the extension of feature classifications from one area to another, and (4) to evaluate the accuracy of the ERTS-1 data in forestry applications.

1.3 GROUND TRUTH

Accurate ground truth acquired from the same time period as the acquisition of the ERTS-1 data was essential. The ground-truth map of the primary area was prepared from aerial photography and was field-checked both on the ground and by helicopter. The ground-truth map of the secondary area was compiled after the classifications were complete to ensure independence from *a priori* knowledge.

The ground truth was acquired from aircraft photography collected in April, October, and November 1972, supplemented by ground checks.

1.4 CLASSIFICATION FEATURES

The 13 classes of features were chosen based on expectations of what the ERTS-1 would be capable of detecting and on the needs of the land manager. A hierarchy of classes was developed and analysis was intended to proceed from the first to the second and third levels. On receipt of the data it was immediately clear that the third level could be achieved, and a fourth level was then regarded as possible.

1.5 ANALYSIS APPROACH

The approach followed in this investigation was to study detection capabilities, then the various conventional imagery interpretation and computer processing methods. The purpose of this was to determine the practicality and validity of using these data in classifying forest and related land-use features. Classification signatures established for the primary area, where 100-percent ground truth was used, were then applied to the secondary area, where ground truth was absent. The resulting maps were evaluated by comparing their classification accuracy with ground truth, and their costs with each other.

1.6 EVALUATION OF DETECTABLE AREA SIZE

The criteria for the minimum size of each feature to be evaluated were established based on expected sensor capabilities and on administrative needs. The administrator must know to what level of detail resources can be mapped

and managed. For the purposes of this study, these criteria were assumed to be 8 hectares (20 acres) for timber types, 15 meters (50 feet) for widths of streams and rights-of-way, and 2 hectares (5 acres) for lakes and rural settlements.

The data studied were (1) the four individual bands of the ERTS multispectral scanner (MSS), (2) composite imagery from the data analysis station/multispectral program, which is enhanced first generation imagery produced by the NASA Lyndon B. Johnson Space Center (JSC) ground data station, (JSC color composite) and (3) simulated color IR imagery from the Goddard Space Flight Center (GSFC).

In summary, the results of this evaluation were as follows. Band 4 (green, .5 to .6 micrometer) was the poorest of the four bands and provided very little information. Band 5 (red, .6 to .7 micrometer) was the best for most features, but was poor for detecting hardwoods and water bodies. A pine stand as small as 4 hectares (10 acres) was detectable when there was good contrast with the surroundings. Band 6 (near-IR, .7 to .8 micrometer) began to show hardwoods and water bodies well, but the detectability of pine stands diminished. Band 7 (IR, .8 to 1.1 micrometers) had similar characteristics to band 6, showing hardwoods best, but was generally slightly poorer than band 6. However, band 7 revealed the smallest features yet seen on ERTS imagery, a group of 1-hectare (2-acre) ponds. The composite made at JSC produced the best detection results and the Goddard color composite rated between bands 6 and 7 (see table 1-I).

TABLE 1-I.- NUMBER OF TARGETS DETECTED

Media	Targets Detected	Percent Detected
JSC-enhanced Composite	21 of 30	70
Band 5	20 of 30	67
Band 6	16 of 30	53
GSFC Composite	15 of 30	50
Band 7	11 of 30	36
Band 4	4 of 30	13

Table 6-I shows the minimum size of forest features which can be detected, with a statement about the contrast with the background which occurred in each case.

1.7 CLASSIFICATION BY CONVENTIONAL METHODS

Conventional interpretation methods were used for four data formats, a single band and three multiband enhancements. The criterion for successful classification was set for the purposes of this study at 80-percent accuracy.

1.7.1 Single-Band Conventional Classification

Since one of the chief breaks in the hierarchy was to distinguish between pines and hardwoods, band 7 was selected because it makes this separation most effectively. The various gray levels were delineated on an enlargement, and these gray levels were then equated with ground truth where possible. Classifications were thus established for these gray levels, and the classes were extended to the rest of the imagery. Only seven classes could be distinguished. The accuracy of the resulting map was determined by overlaying it on the ground-truth map and counting points of agreement versus disagreement by means of a superimposed grid. The accuracy was 62 percent.

1.7.2 Multiband Conventional Methods

The three methods of multiband enhancement available at the JSC, optical combining and two methods of digital combining, are described below.

Optical multiband classification.- This method uses four projectors, which focus the transparencies of the four bands into a single image on a rear projection screen. A red, green, or blue filter was placed over each transparency, and forest features were traced on a color enlargement of this record. Class designations were assigned by comparison with the ground-truth map, and 10 classes of forest features were distinguished. An analysis of the resulting map by the previously used procedure showed 60-percent accuracy.

MCFV classification.- This method employs a scanner which scans three black-and-white transparencies simultaneously, converts the gray levels into digitized codes, and displays the results according to red, green, or blue color values which have been assigned by the operator. Experimentation produced an image which best portrayed the forest features and it was copied photographically. The seven features which could be distinguished were traced on the photographic enlargement and classes were assigned by comparison with the ground-truth map. The analysis showed 69-percent accuracy.

JSC color composite classification.- Using the DAS system, a viewer displays three selected ERTS-1 bands at a time by assigning varying intensities of red, green, or blue. The resulting composite is recorded directly on film and interpreted by standard techniques. Ten classes

of features were distinguished. These features were annotated directly on the film recording and an accuracy of 67 percent was obtained.

1.8 CLASSIFICATION BY COMPUTER METHODS

Two pattern recognition schemes were used in this study; unsupervised classification by a clustering method and supervised classification using a maximum-likelihood method. The criterion for successful classification was again 80-percent accuracy.

1.8.1 Clustering Classification Method

Clustering means that areas in the scene with like spectral values are grouped iteratively into sets called clusters. The clusters are printed out using alphanumeric symbols to represent each cluster. This printout forms a crude map and lines can then be drawn around like symbols to form a classified map. The program can also produce a tape that is compatible with the DAS for display and color film output. The crude printout map produced was converted to a map of the same scale as the ground-truth map. The map had 14 classes and a classification accuracy of 74 percent.

Two methods of clustering were used, a large area containing roughly three times the area under study, and a small area covering little more than the actual area under study. The second method proved to be superior because it was limited to the most pertinent data.

1.8.2 Maximum-Likelihood Classification Method

The computer program used for supervised pattern recognition was LARSYS, a maximum-likelihood technique

developed by the Laboratory for Applications of Remote Sensing (LARS) at Purdue University. In this method the operator supervises the classification by selecting training fields on a preliminary gray map. This tells the computer that when spectral combinations like those in the training field are located, the area should be labeled with the name or symbol of that training field. The classes are printed out as a crude map with symbols representing each class. In the case under consideration, 12 classes were achieved. An accuracy of 70 percent was achieved when compared with ground truth. A tape of the LARSYS classification was then processed on the DAS and recorded on film.

Two other systems using LARSYS classification were tried, the Purdue Remote Terminal and the Earth Resources Interactive Processing System (ERIPS). Both of these systems were found to have practical limitations for use in making forest classifications.

1.9 SIGNATURE EXTENSION

Signature extension involves extrapolating the classification of data from a training area, where it can be supported by ground truth, to a test area where no ground truth exists. The six techniques previously described were applied to the secondary area using only primary area data, and the accuracies achieved were comparable to those achieved in the primary area. This proved that signature extension under these conditions was valid.

1.10 OTHER SIGNIFICANT FINDINGS

Several items of interest to foresters were found in the data incidental to the investigation. An instance was

noted where the effects of forest fires were discovered in the recorded ERTS-1 data. A light ground fire, called a prescribed burn, had been set to clear brush prior to marking timber for sale. The precision with which the effects of this light fire were registered indicates that ERTS data may be used in the future in fire damage assessment to map the perimeter of large fires.

A pine-bark beetle epidemic occurred within the test area during the investigation. Damage was detected indirectly in one case, where approximately 2 hectares of pines were killed in a stand classified as hardwood, although nearly 50 percent of the stand was pine. The LARSYS program classified this portion of the stand as hardwood, cutover. Since this is a supervised program, a beetle-infested training field would have to be included to make such a classification possible.

1.11 ACCURACY AND ANALYSIS

The following table summarizes the accuracy and analysis of this investigation.

TABLE 1-II.- ACCURACY

Technique Primary Area	No. Classes	% Class.	% Acc.	Achievement $\left(\frac{\text{Class.\%+Acc.\%}}{2}\right)$
Single Band	7	50	62	56
Optical Multiband	10	72	60	66
MCFV Multiband	7	50	69	59
DAS Multiband	10	72	67	69
Clustering	14	100	74	87
Maximum Likelihood	12	86	70	78

Higher accuracies were achieved for the key species or major aggregations. Pine, established was classified at 91-percent accuracy, grass at 85 percent and hardwood at 70 percent accuracy by the computer methods. If these classifications had been considered singly, the accuracies would have been much higher. This capability suggests that computer analysis can be used for differentiation of small numbers of major forest classes, such as the first and second levels of the hierarchy shown earlier, with a mean accuracy greater than 80 percent.

1.12 DISCUSSION

The accuracy should be understood in light of its development.

1. Many of the procedures used were new and experimental. This increased the time spent because the procedures were repeated many times under varying conditions until acceptable results were achieved. It was felt that better results could have been achieved

in each case if additional adjustments had been made.

2. A comparison of the results of these procedures is difficult because all the parameters were not held constant. The numbers and kinds of classes changed and different techniques were tried.

3. An adequate picture of ground conditions in this forest test site would require considerably more time than was spent.

4. No precedent was known for comparing classification maps with ground truth and determining their accuracy. As a result the methods used were developed experimentally and may not be adequate.

5. The forest under investigation does not lend itself to exact classification. The timber classes exist in mixtures of species, age, vigor, site, and size classes which blend gradually from one to another.

6. The significance of achieving only 70- to 74-percent accuracy for computer processing should not be minimized. The matching of 14 very complex classes is not easily achieved by any method, and this range of accuracy should therefore be considered high.

1.13 CONCLUSIONS

1. ERTS data can probably be used best in forestry applications if they are used for extensive surveys where broad generalized classes are needed, rather than for intensive surveys where detailed stand conditions must be portrayed.

2. The clustering and maximum-likelihood methods of classification are both efficient, and the difference in accuracies achieved are not significant enough to rank one over the other. However, computer methods as a group can probably be ranked as superior to conventional methods for forest classification.

3. The team was able to classify 15 of the original 19 classes in Level 4 of the hierarchy.

4. In addition to the possibilities of producing the timber inventory maps described above, the ERTS data have the additional potential advantage of providing sequential coverage.

1.14 RECOMMENDATIONS

1. All phases of this investigation should be studied further. Additional evaluations are necessary, especially regarding

   a. seasonal and atmospheric effects,

   b. registration corrections,

   c. temporal correlations,

   d. DAS multiband classification of forest features, and

   e. the removal of striping effects.

2. Greater emphasis should be placed on securing accurate ground-truth maps in future investigations and in determining their accuracy.

3. A better method of statistically evaluating the accuracy of classification maps should be developed.

2.0 INTRODUCTION

2.1 GENERAL

The managers of the National Forests in Texas are in a continual process of inventorying their timber and land resources. This is being accomplished by a combination of low-level aerial surveys and ground observations. The managers of the National Forests in Texas have expressed an interest in finding new techniques for making resource inventories which can be applied over large areas more rapidly and with greater efficiency than the procedures now used.

A team of scientists at the Johnson Space Center, whose specialties include forestry, photogrammetry, imagery interpretation, and automatic data processing, has evaluated the data from ERTS-1 to determine their suitability as a source of data for conducting forest surveys. The team used a portion of the Sam Houston National Forest administered by the National Forests of Texas (see figure 2-1) as a test site for conducting this investigation.

2.2 OBJECTIVES

The general objectives of this investigation were to evaluate the utility of ERTS-1 and underflight data to detect, identify, and determine the areal extent of features related to a forestry application in a limited study area within the Houston Area Test Site.

The specific objectives of the study were (1) to determine the size of forest features which can be detected in the data under different conditions; (2) to determine the

NASA S-73-18483

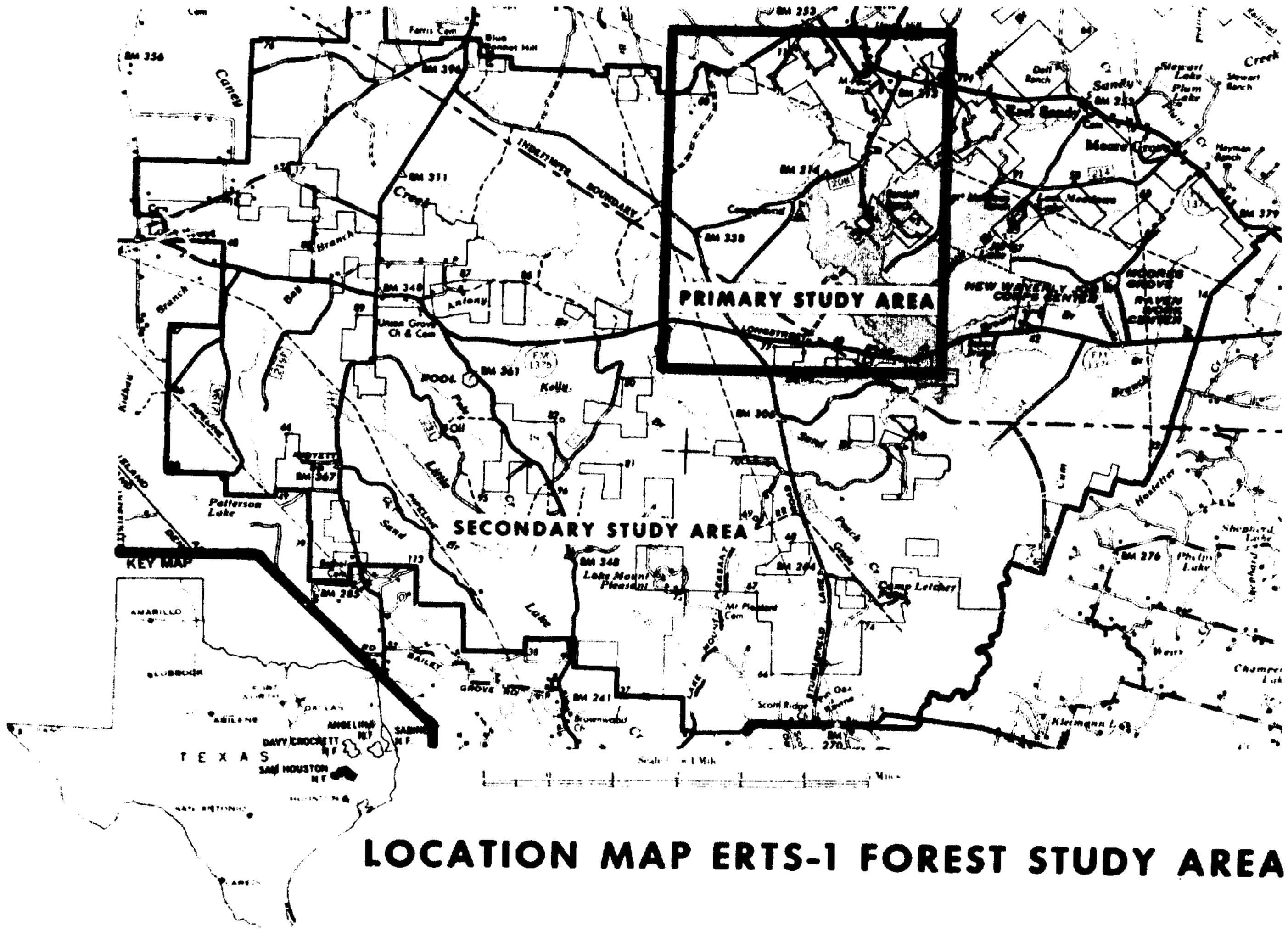


Figure 2-1.- Location map of ERTS-1 forestry study area.

suitability of the data for making classification maps showing timber species, stand conditions, and related land uses by both conventional and computer-processing methods; (3) to test the extension of classification data from one area to another; and (4) to evaluate the accuracy of these operations.

3.0 STUDY AREA, FEATURES, AND CRITERIA

3.1 USER REQUIREMENTS

This investigation was part of a long-range study conducted to determine the extent to which the ERTS and other remote sensing systems could contribute information of interest to forest managers. Forest administrators need up-to-date maps covering the resources they manage. They need, and may possibly secure from the ERTS, timber stand and condition class maps, maps showing changes in cutover and regeneration areas, and maps showing improvements such as roads, bridges, and water bodies. Resources less likely to be monitored by ERTS are soils, range conditions, and wildlife habitats.

3.2 STUDY AREA SELECTION

A small portion of the Sam Houston National Forest (SHNF) and vicinity was selected as the primary ERTS-1 forest study area to take advantage of the remote sensing and ground work done by the Earth Observations Division personnel and the ground-truth data collected by the personnel of the National Forests in Texas. The area was the north central portion of the Raven Ranger District (shown in figure 2-1), which is approximately 6,000 hectares (15,000 acres).

A secondary study area was used as a test site for the extension of information obtained in the primary area. It was composed of the western part of the Raven Ranger District, SHNF. Since the conditions surrounding the primary study area were similar to those within the primary area, they

would provide validity for feature signature extension. This secondary study area was a continuous block of land roughly 30,000 hectares (82,200 acres) in size (also shown in figure 2-1).

3.3 FEATURES

The features studied were surface conditions as defined by vegetative cover types and condition classes or land-use classifications. Table 3-I provides the hierarchy of features.

TABLE 3-I.- HIERARCHY OF FOREST FEATURES

Level 1	Level 2	Level 3	Level 4
Forest	Standing Timber	Pine, established Hardwood, established	Pine, established Hardwood, established
	Cutover Timber	Pine, cutover Pine, site prepared Pine, regenerated Hardwood, cutover	Pine, cutover Pine, site prepared Pine, site prepared/ vegetation Pine, regenerated Hardwood, cutover
Nonforest	Water	Streams and lakes Impoundments	Streams and lakes Impoundments
	Other Land-use Classes	Brush Cultivated Grass Rights-of-way Rural settlement	Weeds Brush Cultivated Rights-of-way, dirt roads, or utility lines Grass Rights-of-way, paved roads Cut and bulldozed Bare soil Rural settlement

The 13 classes of features (Level 3) chosen for analysis were based on what ERTS was expected to be capable of detecting and on the needs of the land manager. The hierarchy of classes in table 3-I was developed with the intention that analysis would start with the first level and proceed to the second and third levels (ref. 1). Inspection of the first imagery showed that the third level could be achieved immediately. During the first analysis task, feature detection, all 13 Level-3 features could be detected.[1] However, during the study additional classes were distinguished. A fourth level was then required, and these classes are listed with a definition of each.

1. Pine, Established

   This class consists of stands with a canopy containing 50 percent or more loblolly or shortleaf pine. The canopies of these stands must contain a majority of trees which are greater than sapling size. The upper limit of saplings at breast height (4-1/2 feet above ground) is 12.7 cm (5 inches) in diameter.

2. Pine, Cutover

   An area where the pine timber has been cut but where the ground has not been prepared for replanting. Such areas usually contain a large amount of logging debris, numerous pine and hardwood trees too small to be cut, a few larger cull trees (usually hardwood),

---

[1]The 13 classes were not all detected with equal speed or accuracy. Those with good contrast with background gave good results; those with poor contrast gave poor results (refer to paragraph 6.1).

and the original ground cover of grass and shrubs. Pines, to be detected by remote sensing, must be in the majority among the remaining trees.

3. Pine, Site Prepared

   An area of cutover pine where the ground has been prepared for replanting by knocking down and breaking up the remaining logging debris and ground cover with heavy mechanical equipment. Mineral soil is exposed over much of the area and the spectral return is a combination of bare soil and logging debris. Narrow strips of hardwoods along streams and small islands of hardwoods are frequently left as cover for wildlife.

4. Pine, Site Prepared/Vegetated

   This is a special classification not recognized by forest managers, which was created because of the limitations of small-scale imagery. It represents the intermediate stage between site preparation and regeneration. During this period grass, weeds, shrubs, and possibly planted pine seedlings may occupy the site, but the pine seedlings are too small to be detected. Some bare soil and logging debris are also visible, making it a blend of spectral values which provides a signature.

5. Pine, Regenerated

   These stands consist of seedling- and sapling-sized pines which have formed a canopy dense enough to be visible on small-scale imagery. In this study they are usually plantations which have been created by man following cutting, but in some instances they consist of naturally regenerated stands.

6. Hardwood, Established

   This class consists of stands of tree species where the canopy is 50 percent or more deciduous. The majority of the trees in the canopy must be greater than sapling size.

7. Hardwood, Cutover

   An area which was previously covered by hardwood or pine timber from which all the merchantable timber has been removed, with the majority of remaining saplings and poles being deciduous tree species. This places the emphasis on the class at the time of viewing, rather than upon its historical classification, which may not be known.

8. Grass

   An area occupied either by grass or herbaceous plants, where grass predominates. This may be natural in origin, planted, or otherwise improved pasture, which has been in existence long enough to have lost the parallel lines of cultivation.

9. Weeds

   An area of grass and herbaceous plants where the herbaceous plants predominate.

10. Cultivated

    An area of land either prepared for or occupied by subsistence or commercial crops or introduced forage plants, where the rectangular outlines and parallel interior lines of cultivation are still visible.

11. Streams and Lakes

    A natural body of water.

12. Impoundments

    A body of water formed by artificial entrapment. This class was separated from streams and lakes because it usually contains shallow water, which is lighter in tone than the deeper natural bodies of water.

13. Rights-of-Way, Dirt Roads, or Utility Lines

    Cleared strips of land which are occupied by a utility transmission line, a dirt road, or both.

14. Rights-of-Way, Paved Roads

    Cleared strips occupied by a hard-surfaced paved road.

15. Rural Settlements

    Areas containing widely spaced residential construction and farm buildings.

16. Cut and Bulldozed

    A special classification created to cover the land clearing in the bottom of Lake Conroe, where the trees and shrubs had been removed and burned but the top soil and sod remained.

17. Bare Soil

    An area where the top soil and sod has been removed, exposing the mineral soil.

18. Brush

    An area of land covered by shrubs of species which remain shrubs throughout their life cycle. This

class was deleted because of its limited occurrence in the study area.

3.4 CRITERIA

The criteria established for the minimum size of each feature to be evaluated were based on the expected ERTS-1 performance and assumed administrative needs. These criteria were 8 hectares (20 acres) for timber types, 15 meters (50 feet) for width of streams and rights-of-way, and 2 hectares (5 acres) for water bodies and rural settlements. The criterion for successful feature classification was set at 80-percent accuracy. This accuracy target was set by the managers of the National Forests in Texas as one which seemed appropriate for an investigation of this type, but there was no operational precedent for its selection.

4.0 DATA UTILIZATION

4.1 ERTS-1 DATA

System-corrected multispectral scanner (MSS) tapes and imagery from the August 29, 1972, mission (frame 1037-16244) were used for most of the classification work. A followup study was done using data from the November 27, 1972 mission (frame 1127-16253) over the Houston area.

4.2 AIRCRAFT DATA

The ground-truth map of the primary study area was prepared at a scale of 1:24,000 from aircraft photography flown April 1972, at an altitude of 18 kilometers (60,000 ft). The data were acquired using a 150-mm (6-in) focal length RC-8 camera and color infrared (CIR) film. Underflight photography was flown in August 1972, within a day of the ERTS-1 coverage, but because of high cloud cover the team was forced to rely on the earlier photography. The results were checked against the aircraft coverage of October 2 and November 7, 1972. The few changes which occurred in the intervening months (May through October) caused problems. The primary problem was determining whether cutting the site preparation operations occurred before or after the ERTS-1 coverage. The film was viewed under a stereomicroscope, classifications were drawn on an overlay, and the classification map, shown in figure 4-1, was compiled.

The ground-truth map for the secondary study area was compiled at a scale of 1:63,360 from aerial photography from the three coverages described. This map was compiled

NASA S-73-28187

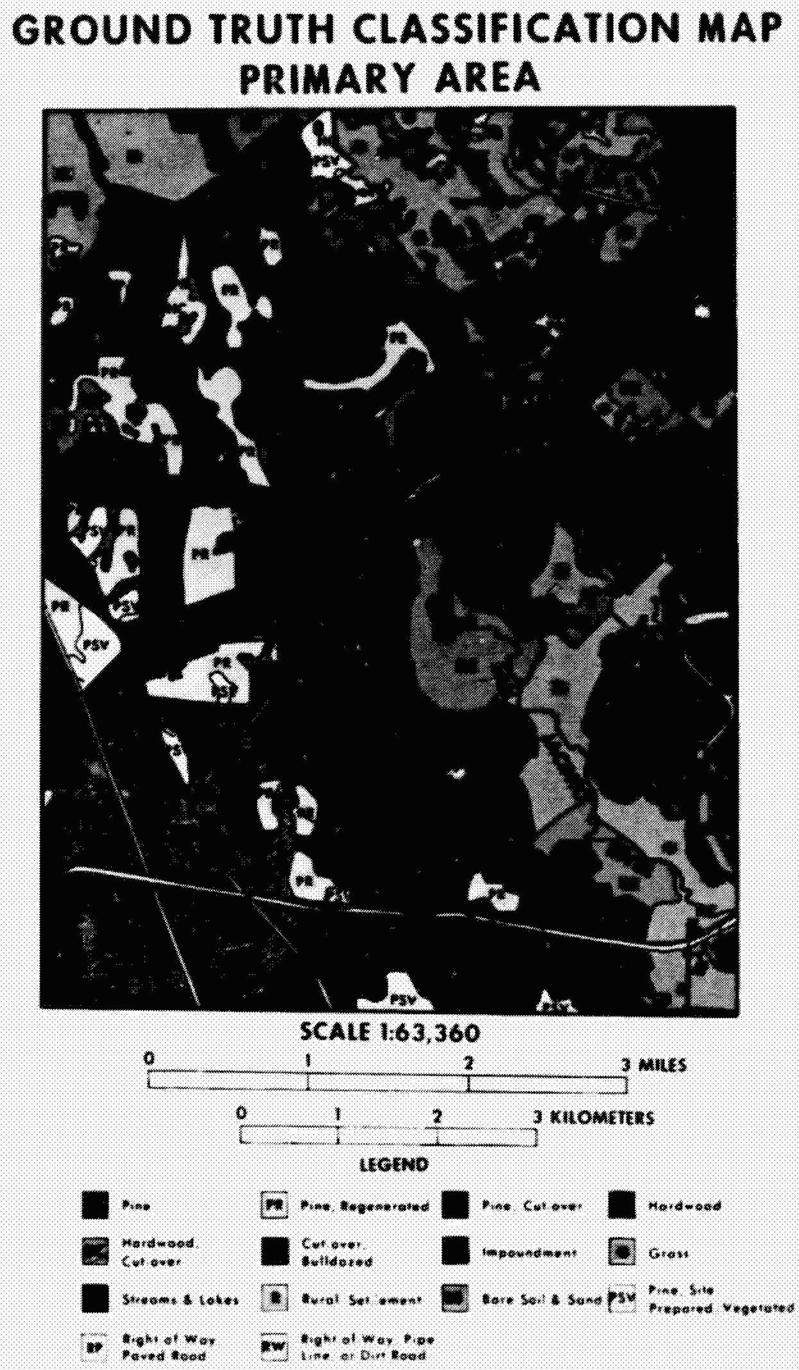


Figure 4-1.- Ground-truth classification map of primary area.

after the classifications of the secondary area were completed to ensure that these classifications were independent of the analysis.

4.3 GROUND SURVEY

The ground-truth map was spot checked in the field by visiting points on the ground where questions had arisen during the interpretation of the photography. The entire area was then checked by an observer in a helicopter at a low altitude.

Experience has shown that accurate ground truth from the same time period as the acquisition of the ERTS-1 data is essential. Ground truth is defined as a broad classification of timber species and conditions and related features in map and area listing forms. Ground truth is the only reliable means of determining the performance of the ERTS-1 sensor data for their use in forest classification surveys, because it provides a standard by which the data can be measured.

5.0 ANALYSIS APPROACH

Various conventional imagery interpretation and computerized processing methods were used to determine the practicality and validity of each in classifying forest and related land-use features. The resulting maps were evaluated by comparing their classification accuracy with ground truth. Their cost was evaluated by comparing their unit cost with each other.

5.1 SENSOR DETECTION EVALUATION

This study was tied directly to user requirements. Forest managers routinely map features according to the minimum size they plan to manage as a workable unit. Therefore, it is important for them to know the minimum sizes they can expect from the analysis of ERTS-1 data.

The minimum area of forestry features that could be detected by the sensors was determined. This was done by locating examples of each of the features listed in Level 3 on underflight photography and then on the ERTS-1 imagery. These features were then evaluated using the criteria established, to see how size affected detection.

5.2 CONVENTIONAL IMAGERY INTERPRETATION PROCESSING

Four methods were available for performing a conventional analysis: (1) single-band classification, (2) optical multiband classification, (3) MCFV classification, and (4) JSC color composite classification.

5.2.1 Single-Band Classification

Very little previous work has been done by foresters on ultra-small-scale photography from high altitude, other than that done by Aldrich (ref. 2), whose work was reviewed to provide a background for this study.

One of the important differentiations in the hierarchy of feature classification is that of pines from hardwoods. The MSS band 7 was selected because it most easily separates these species. The various gray levels were delineated on an enlargement by ocular estimation of the various levels. These gray levels were then equated with ground truth where possible. Classifications were established for these gray levels, and the classifications were extended to the balance of the imagery.

5.2.2 Multiband Classification

A number of multiband methods of classification have been used to exploit their capability to enhance selected features. The three methods available for use were optical combination and two digital methods, and these are described below.

Optical multiband classification.– Several instruments for optically combining the four electron beam recorder (EBR) recordings of the ERTS-1 imagery were tried, and the most satisfactory results came from the multispectral additive color viewer. The best portrayal of forest features was produced when bands 4 and 5 had a green filter, band 6 had a blue filter, and band 7 had a red filter. Forest features were then traced on a color enlargement of this record, and

class designations were assigned by comparison with the ground-truth map.

MCFV classification.- A second multiband viewer offered another method of combining ERTS-1 imagery. This method employed a scanner which scanned three images simultaneously, converted the gray levels into digitized codes, and displayed the results according to red, green, and blue color values which were assigned by the operator. Experimentation produced an image which best portrayed the forest features and it was copied photographically. The features were then traced on the photographic enlargement and classes were assigned by comparison with the ground-truth map.

JSC color composite classification.— The data analysis station (DAS) provided the third method of multiband viewing. Using the MOPS program this viewer used the ERTS-1 tapes and displayed three ERTS-1 bands at a time by assigning various intensities of red, green, or blue to the three bands selected. Bands 5, 6, and 7 gave the best performance for forest features, and these results were recorded directly on film.

5.3 CLASSIFICATION BY COMPUTER METHODS

The two pattern recognition schemes used in this study were unsupervised classification by a clustering method and supervised classification using maximum-likelihood methods.

5.3.1 Clustering Method

Clustering is a computer program which groups similar spectral values iteratively into sets called clusters. Clustering has been used extensively for determining boundaries, the homogeneity of data sets, and to support the LARSYS pattern recognition program for selecting training and/or test fields. Clustering has been used most recently as an unsupervised pattern recognition scheme. However, when using clustering as an unsupervised classifier it is very difficult to determine the number of iterations required. The only criterion for deciding how many iterations were necessary was what best matched the ground-truth feature patterns on the gray-map printout. The analyst examined each iteration until he found the one with feature patterns which best matched the ground truth. This is considered to be a serious weakness of the system.

If the user has some knowledge of the data or the area that is to be clustered, a set of mean values (one for each cluster) may be entered into the program to help initiate it and to reduce processing time. Also, by setting up certain threshold values for the means, standard deviations, and number of clusters, the user can control the program's function to some degree. This may reduce the number of iterations somewhat, but several iterations are still required. The clusters are printed out using a symbol from an alphanumeric code to represent each cluster. This printout forms a crude map and lines can then be drawn around similar symbols to form a classification map. The program is also capable of producing a tape that is compatible with the DAS for display and color film output.

Clustering was used as a classifier for both large and small areas. In the first method a large area approximately three times the area contained in the primary study area was classified. Because of the problems which occurred because the area was so large, clustering was then limited to an area not much larger than the primary study area.

5.3.2 Maximum-Likelihood Method

The computer program for supervised pattern recognition for earth resources is the maximum-likelihood scheme LARSYS, the Purdue University program. Three versions of this program were used to evaluate ERTS MSS data: (1) the Purdue version of LARSYS on the IBM 360/67 computer accessed by a remote terminal, (2) the modification of LARSYS for the computers at JSC, and (3) the Earth Resources Interactive Processing System (ERIPS) at JSC.

Purdue terminal classification.– The Purdue terminal has proved to be an excellent means for securing classifications. Although the time required to send an input tape to Purdue and to receive an output tape is a drawback, this is balanced by the speed with which a classification can be made using the terminal once the tape is at Purdue. The team used the terminal minimally, however, because the tape sent to Purdue contained errors. The tape was used to obtain a fast reading on what certain combinations of training fields would produce, and these results were then used on the 1108 computer.

LARSYS classification.– The LARSYS supervised classification program was used on the JSC computers. A methodology similar to that used in clustering was applied to LARSYS. That is, in using the LARSYS classifier numerous computer runs were made before the best combination of classes and training fields was found.

ERIPS classification system.– The Earth Resources Interactive Processing System (ERIPS) theoretically has great potential in analyzing data from various imaging sensors. It was tried so that all available systems could be appraised.

5.4 SIGNATURE EXTENSION

Signature extension involves extrapolating the classification of data from a training area where it is supported by ground truth to another area where no ground truth exists. The term is used here in a limited sense because extensions at different time intervals and atmospheric conditions were not attempted. An analysis was conducted to determine if the signatures found in the primary study area could be applied to the surrounding secondary study area. This made the primary area the training area and the secondary area the test area. Ground truth was necessary to determine the efficiency of such a procedure even in the extended or secondary area. However, the secondary area ground truth (shown in figure 6-1) was not available to the interpreter until accuracy was computed.

5.4.1 Conventional Classification Extension

Single-band classification extension.– The features of the secondary area were classified by using the results of

the primary area classification as a guide. Certain tonal ranges were related to classes, and these relationships were extended into the secondary area.

Multiband classification extension.— Information extension by multiband methods was conducted similar to the single-band method. Color-enhancement enlargements of the secondary area were made under parameters identical to those for the primary area. Since various color values had been equated to classes in the primary area, these values were used as signatures in the extension of information.

5.4.2 Computer Classification

In the primary area 100-percent ground truth was necessary to equate clusters and cluster combinations to classes. Since a much larger area had been included in the computer printout which covered most of the secondary area, the same signatures were extended into this larger area without ground truth. In the LARSYS method, training fields were selected only in the primary area. The computer printout continued into the secondary area and the resulting classes were not supported by additional training fields. These classes were used, although they diminished in validity on the periphery of the secondary area where vegetation changes began to occur.

5.5 ACCURACY EVALUATION

The approach used in accuracy evaluation was to overlay each classification map on the ground-truth map and the points of class agreement were statistically counted.

A 1/4-inch grid (approximately 500 feet) was overlaid on the two maps, and the points of agreement versus disagreement were counted. A grid intersection was counted wherever it fell on a point of agreement between the two maps. This permitted the percentage of agreement with ground truth to be determined. The total of agreement points, divided by the total of all points, resulted in the percent of accuracy. This focuses attention on accurate boundary definition, and is not related to classification performance (paragraph 1.4), which is a measure of homogeneity within a class.

The accuracy of the classification map was also assessed by determining the amount of area of each class when compared with ground truth. The total area of each class was listed, and the difference between these values and the ground-truth area values was computed. The total of the differences between the areas was the disagreement between maps, and this permitted an accuracy percentage to be computed. This was a less precise method of calculation because it considered only the total area of each class and not the position of the classes. Presumably all the classes could be out of position and yet could produce 100-percent accuracy in area quantities.

6.0 ANALYSIS RESULTS

6.1 EVALUATION OF DETECTABLE AREA RESULTS

The original plan was to determine how differences in the atmosphere, season of the year, and feature size affected the detection of the third level of the hierarchy. Differing atmospheric conditions would also affect the size of features that could be detected. The season of year would also be a factor. For example, in winter when the hardwoods have lost their leaves the line between hardwoods and pines would be more sharply delineated and smaller features could be detected than at other times of the year. However, time limitations restricted the study to feature size detection for only one set of data.

The four bands of the multispectral scanner were reduced to black-and-white imagery and studied, as well as a JSC color composite and a GSFC election beam recorder (EBR) type-C composite.

The minimum size of forest features which could be detected are listed in table 6-I. Table 6-II shows the number of the targets and the sizes which were chosen for study, and table 6-III shows the success with which representative classes were detected.

The results of this evaluation were as follows. Band 4 (green, .5 to .6 micrometer) was the poorest of the four bands and provided very little information. Band 5 (red, .6 to .7 micrometer) was the best for most features, but was poor for detecting hardwoods and water bodies. A pine stand

TABLE 6-I.- MINIMUM SIZE OF FOREST FEATURES

	Smallest Size Detected (hectares)	Band or Composite	Contrast With Background
Pine, Established	4	5,7, JSC Comp	Good, Fair, Fair
Pine, Regenerated	8	5, JSC Comp	Poor, Fair
Pine, Site Prepared/ Vegetated	8	5	Fair
Pine, Cutover	24	5, JSC Comp	Good, Fair
Hardwood, Established	[a]12	JSC Comp	Fair
Brush	4	5,7	Fair, Poor
Grass	6	JSC Comp	Good
Cultivated	4	5, JSC Comp	Good, Fair
Streams and Lakes	2	5,7, JSC Comp	Poor, Good, Poor
Impoundments	1	JSC Comp	Poor
Rights-of-way	28 meters	5, JSC Comp	Fair, Fair
Rural Settlement	40	5	Poor

[a]No smaller hardwood stands with good contrast with background could be found in the study area. If they could have been located, they would have been detected most readily on band 7. The 12-hectare stand selected blended gradually into the surrounding pines which was not detectable on band 7, but was detectable on the JSC composite.

TABLE 6-II.- TYPES OF DETECTION TARGETS

Detection Targets	No. of Targets[a]	Sizes (hectares)
Pine, Established	5	4,10,12,20,75
Pine, Regenerated	3	8,12,42
Pine, Site Prepared/ Vegetated	4	8,10,16,24
Pine, Cutover	1	24
Hardwood, Established	2	12,42
Brush	2	4,24
Grass	3	6,12,32
Cultivated	2	4,8
Streams and Lakes	2	2,12
Impoundments	2	1,4
Rights-of-way	2	14 meters, 28 meters
Rural Settlements	2	13,40
	30	

[a]The number of targets is generally proportional to the area in each class.

TABLE 6-III.- PERCENT OF TARGETS DETECTED

	Pine, Estab.	Pine, Regen.	Pine, Site Prep./Veg.	Grass	All Targets
JSC DAS Composite	80	100	75	100	70
Band 5	60	67	100	67	67
Band 6	20	67	50	100	53
Goddard Type-C EBR Composite	40	67	75	67	50
Band 7	40	67	25	33	27
Band 4	0	33	0	33	13

as small as 4 hectares (10 acres) was detectable when there was good contrast with the surroundings. In band 6 (near-IR, .7 to .8 micrometer) hardwoods began to appear and water bodies stood out well, but pine stands diminished in detectability. Band 7 (IR, .8 to 1.1 micrometers) had characteristics similar to band 6, showing hardwoods best, but was generally slightly poorer. However, band 7 revealed some of the smallest features yet seen on ERTS-1 imagery, a group of 1-hectare (2-acre) ponds. The composite made at JSC produced the best detection results of all the media examined; the Goddard color composite rated between bands 6 and 7 (table 6-IV). These results were significant because they show the limitations imposed on conventional photointerpretation methods.

TABLE 6-IV.- NUMBER OF TARGETS DETECTED

Media	Targets Detected	Percent Detected
JSC Color Composite	21 of 30	70
Band 5	20 of 30	60
Band 6	16 of 30	53
Goddard Type-C EBR Composite	15 of 30	50
Band 7	11 of 30	36
Band 4	4 of 30	13

These data indicated that forest stands in the test area could be detected in sizes ranging from 4 to 12 hectares (10 to 30 acres). However, this study was conducted with known targets, and there is a considerable difference between the detection of a known target and the identification of an unknown target. Studies by Heyning (ref. 3) established a resolution ratio of 4:7:10 between the size requirements of detection, recognition, and identification. There is a direct relationship between resolution and target size if only spatial considerations are involved. Therefore, using this ratio, recognizing an unknown stand and separating it from others should require areas ranging from 7 to 21 hectares (17 to 52 acres). The positive identification of a stand would require areas from 10 to 30 hectares (25 to 74 acres). However, since spectral values are also involved, the desired sizes are between these extremes. This helps to explain why the conventional methods, which depend on human detection, recognition, and identification from black-and-white imagery, produce relatively poor results, and why

computer methods, which use spectral differences, give better results.

6.2 RESULTS OF CLASSIFICATION BY CONVENTIONAL METHODS

Testing conventional classification methods was one of the objectives of this study, and to adequately test the data all interpretative techniques were employed. The conventional methods were generally time-proven techniques and provided a comparison with computer techniques. However, only one of the four conventional methods investigated proved to be adequate for the ERTS-1 data.

6.2.1 Single-Band Classification Results

Only seven classes could be distinguished. The accuracy of the resulting map (shown in figure 6-1) was determined by superimposing it over the ground-truth map. The accuracy, determined by statistically counting the points of classification agreement, was 66 percent. The accuracy by the area computation method was 58 percent. The mean of the two percentages was 62 percent.

Discrimination was generally poor between the forest features found in this study area when unenhanced imagery was used. This indicated that the single-band method offered very little for forest classification mapping, and is not recommended for future work.

6.2.2 Multiband Classification Results

Optical multiband classification results.— Ten classes of forest features were distinguished by this method. The resulting map (figure 6-2) was analyzed by the previously

NASA S-73-25527

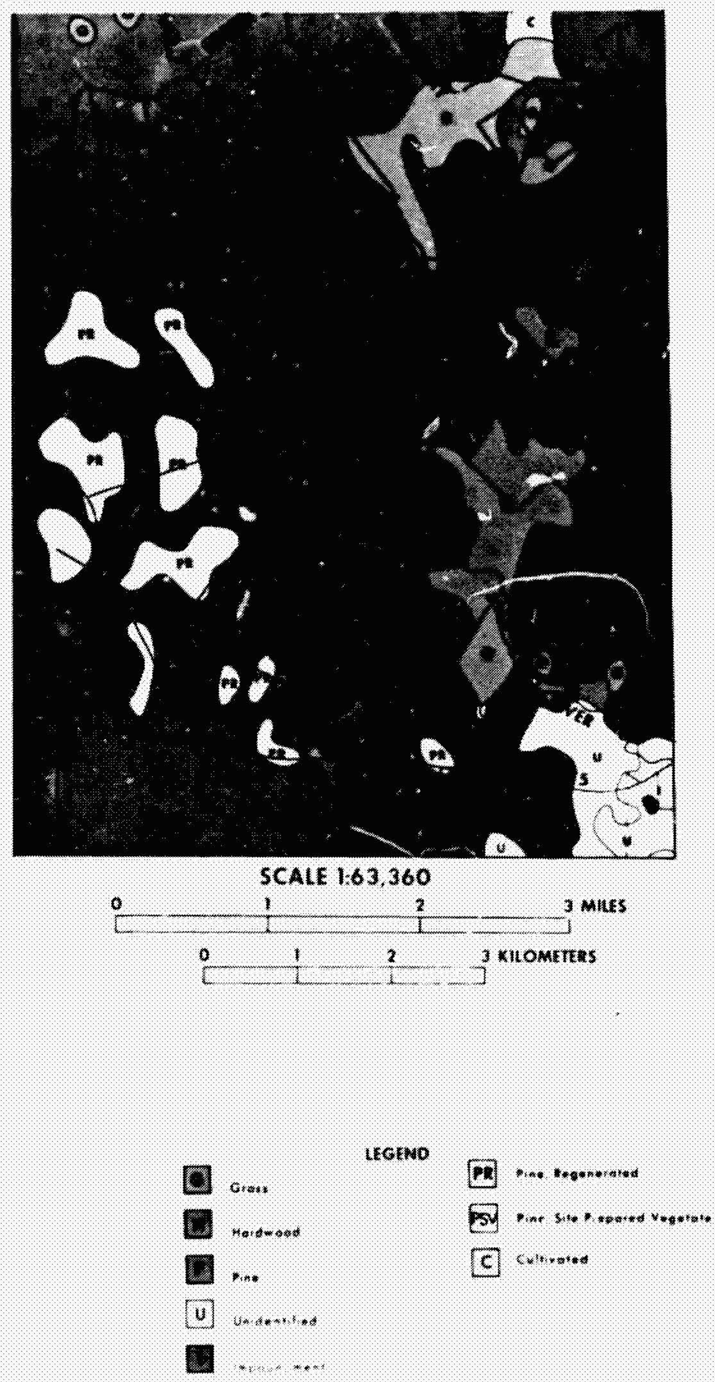


Figure 6-1.- Single-band conventional classification map of primary area.

NASA S-73-28195

MULTISPECTRAL CLASSIFICATION MAP
PRIMARY AREA

LEGEND

- Hardwood
- Pine
- PS Pine, Site Prepared
- PR Pine, Regenated
- Pine, Cutover
- RP Right Of Way Paved Roads
- RW Right Of Way Pipeline & Dirtroad
- Baresoil & Sand
- Grass
- Hardwood, Cutover

Figure 6-2.- Optical multiband classification map of primary area.

used procedures, and a mean accuracy of 60 percent was obtained.

During the investigation the additive process of rear-view screen projection was determined probably not to be the best method of optical multiband analysis. The diazo method of combining various bands (using a different sheet of colored diazo film for each band) is quicker, cheaper, and much easier to manipulate than an additive color viewer. For this reason the use of an additive viewer is not recommended for subsequent work. The recommended procedure is to print a colored diazo transparency of each of the four bands of imagery using a minimum of red, green, and blue; additional colors such as yellow and purple may also be used. This provides a large number of combinations in which the transparencies can be assembled to produce the best enhancement of features.

MCFV classification results.- Seven forest features were classified by this method, and a mean accuracy of 69 percent resulted. The map is shown in figure 6-3.

The use of the MCFV for analyzing ERTS-1 data is regarded as inefficient because it must redigitize data which have been reduced to an image for gray-level slicing. To make efficient use of the ERTS-1 data, the ERTS-1 tapes should be input, rather than the reconstituted images. A recording device which transfers the digitized data directly onto film, similar to the DAS, is recommended.

NASA S-73-28188

Figure 6-3.- MCFV classification map of primary area.

To make the best use of the instrument in its present form, the film input should have generally low density and contrast to produce the best results. The film should not be enlarged beyond the 9-by 9-inch format.

JSC color composite classification results.– Ten classes of features were distinguished by this method of classification. These classes were annotated directly onto the DAS recording as shown in figure 6-4 where the distortions of skew and scale remain. These distortions are caused by the earth's rotation effect and by the inequalities of horizontal and vertical scales in the printer, which can be removed. This method produced an accuracy of 67 percent.

Following the compilation of the data, further experimentation showed that considerable improvement in discrimination of features was possible. Several combinations of instrument settings showed a variation in pine stands which needs further study. This variation could represent species or stand condition differences, but time and limited resources prevented the team from investigating this possibility further. The JSC color composite method is considered to have great promise and much more experimentation is warranted. An example of improved JSC color composite imagery is shown in figure 6-5.

6.3 RESULTS OF CLASSIFICATION BY COMPUTER METHODS

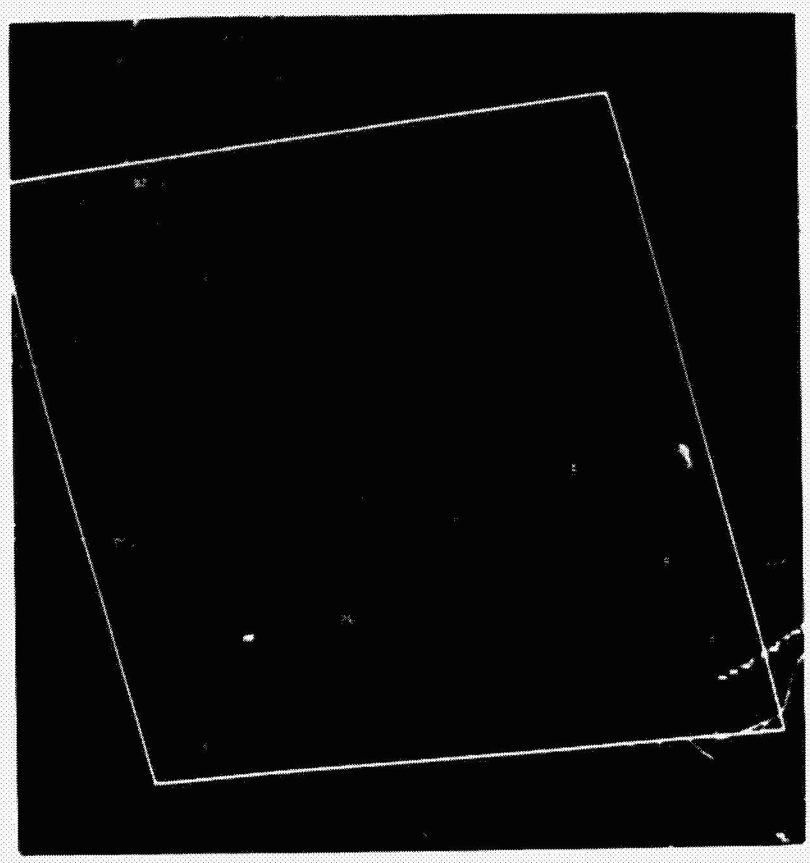
6.3.1 Clustering Results

Large-area cluster classification results.– In the study conducted by the team, the ERTS-1 MSS tape covering the primary area was used. The tape was edited to an area 300

NASA S-73-1124

FOREST CLASSIFICATION OF SHNF PRIMARY STUDY AREA

FROM DAS/MOPS COMPOSITE



LEGEND

P	Pine, Established
PP	Pine, Regenerated
PS	Pine, Site Prepared
PSV	Pine, Site Prepared/Vegetated
H	Hardwood
HC	Hardwood, Cutover
G	Grass
I	Impoundments
BS	Bare Soil
CB	Cutover & Bulldozed

Figure 6-4.- JSC color composite classification map of primary area.

NASA S-73-38512

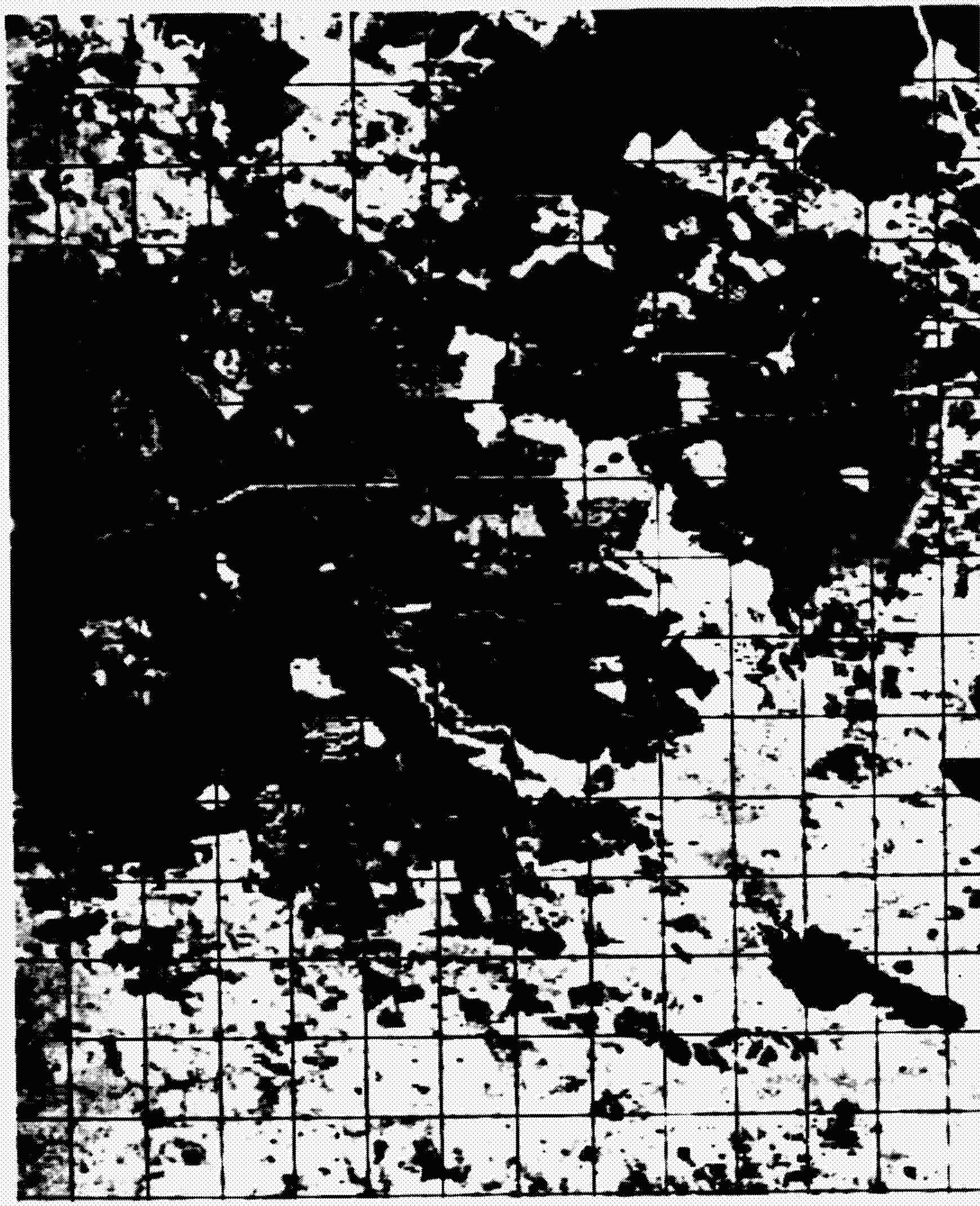


Figure 6-5.- Improved JSC color composite of the study area. (The superimposed grid is not relevant to the problem.)

columns of pixels wide by 200 lines of pixels deep, which is about three times the size of the study area. The STDMAX (maximum standard deviation) and DLMIN (minimum distance between cluster means), the two main control parameters, had to first be determined. A recent publication by Kan and Holley (ref. 4) suggests that a STDMAX of 4.5 and a DLMIN of 3.2 be set. On the first attempt the STDMAX proved to be too large, since by using 4.5, the complete area was clustered into only one category. A number of runs were made reducing STDMAX by 0.5 time until a value of 2.0 was reached.

In figure 6-6 cluster number 1, which represents pine, established, had standard deviations of 1.12, 0.77, 1.95, and 1.20 for channels 1, 2, 3 and 4,[2] respectively. The standard deviation of 1.95 in channel 6 is just below the 2.0 value, and to lower the STDMAX to less than 1.95 would cause the pine cluster to split into two or more categories. Cluster 7, which represents hardwood, established, has standard deviations of 1.06, 0.99, 1.79, and 1.15 for the four channels, respectively. Cluster 1 has a total of 36,037 points, which represent about 60 percent of the total numbers of points, and cluster 7 has a total of 7,397 points, or 12 percent of the total. Relating these percentages to the ground-truth values (the ground truth was 54 percent for pine and 13 percent for hardwood) indicates that the clustering technique for these two features was well within the expected tolerances.

---

[2]According to the GSFC "Data Users Handbook," the terms "channels 1,2,3, and 4" equate to "bands 4,5,6 and 7". The computer programs use the channel numbers exclusively, and the multispectral procedures use the band numbers.

NASA S-73-28214

Figure 6-6.- Computer printout of cluster classification of primary area.

Seventeen clusters of a total of 26 had a standard deviation of less than 2.0. The other clusters, with a standard deviation larger than 2.0, had only a small number of points representing these clusters, with the exception of cluster 25 which had 1,792 points. This cluster and cluster 4 should and would have been combined into one cluster if the program had been allowed to continue another iteration, because the distance (DLMIN) between the two was only 2.8, which was less than the 3.2 used for a threshold value. This eventually produced a run with a maximum of 30 clusters, which then recombined to 24 clusters.[3]

In figure 6-6, which illustrates the printout of the 24-cluster map, the symbol 1 is most prevalent, and forms a pattern which most closely approximates the pine, established, classification in the ground-truth map. The pattern produced by the symbol 7 appears most often where hardwood, established, is found, and there the coincidence of single symbols with classification features ends. It was then observed that certain combinations of symbols seemed to fall into patterns which represented other classifications. For example, wherever there was a pine, site prepared class, the symbols 2, 4, 7, and 8 appeared together, with the 2's predominating. Using this methodology the team devised a code of individual symbols and combinations of symbols which represented all of the ground-truth classes and provided a set of signatures for each class. Hoffer, et al. (ref. 5) used this technique of cluster combinations in a recent study to classify fallow fields, which had the wide range of spectral values similar to cutover and regenerated forest classes.

---

[3]The computer program splits the clusters until they reach a predetermined maximum and then recombines them until an equilibrium is reached.

The signatures and ground-truth classes did not correspond completely, but the signatures were followed in compiling the cluster map, regardless of possible disagreement with ground truth. A few additional classes appeared in the clustering, since the clustering split rights-of-way into two groups, those containing dirt roads or utility lines and those containing paved roads. It also split hardwood, cutover into hardwood, cutover and cut and bulldozed, the latter a class in a river bottom which was being cleared for a lake floor. Areas of bare soil were recognized where bulldozing had cleared away all surface vegetation and top soil, as well as what appeared to be cultivated land in the grass classification. This was particularly remarkable because the cultivation patterns were so subtle that they were not visible in the aircraft photography, and had not been included in the ground-truth map. Ground inspection showed that the cultivation was "improved pastures," and that fertilizing and liming of the soil produced a different spectral return than the unimproved pastures.

A total of 14 classes were recognized. The resulting printout map was converted to the base map scale and the line drawing shown in figure 6-7 was produced. The agreement computation made by comparison with ground truth was 74 percent. The most prevalent error was to classify pine, regenerated, as hardwood, cutover. This was an understandable classification because the mature pine had been removed from the regeneration areas, and the ground cover was a mixture of small planted pines and dense hardwood reproduction.

NASA S-73-28189

CLUSTER CLASSIFICATION MAP
PRIMARY AREA

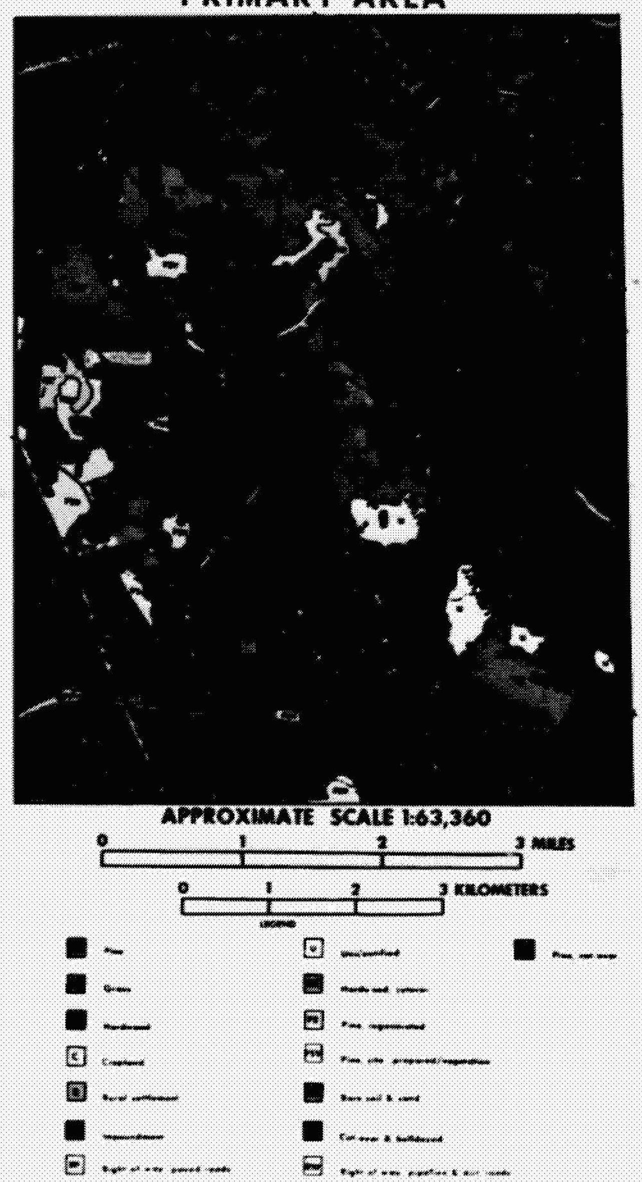


Figure 6-7.- Classification map of primary area derived from clustering.

The results from the clustering program were also recorded on the DAS, and this is shown in figure 6-8. These classifications were not as accurate as those on the map made directly from a machine printout because the printout method made possible a combination of the various symbols which constituted a signature. A given color may be used with only one symbol at a time with the DAS recording, and combinations are not possible if a symbol is to appear in more than one class. As a result, the DAS recording was only about 60-percent accurate. The illustration shows the film output with its inherent distortions of skew and unequal horizontal and vertical scales which still need to be corrected.

The statistics of this cluster run are shown in tables A-III and A-IV, of the appendix. The standard deviations are quite low and uniform and there is adequate spacing between clusters. This indicates that the features were classified efficiently.

Small-area cluster classification.— A final clustering run was made to determine if a reduced amount of data would improve the clustering results. The number of columns of pixels were reduced from 300 to 170 and the lines of pixels from 200 to 125. This reduced the area to very little more than the primary study area. The other constraints remained the same, with the exception that processing was stopped at the 12*th* iteration and at 21 clusters. This produced a printout which was very close to the ground truth.

Nine combinations of the symbols were made to form classes which required the use of each symbol only once.

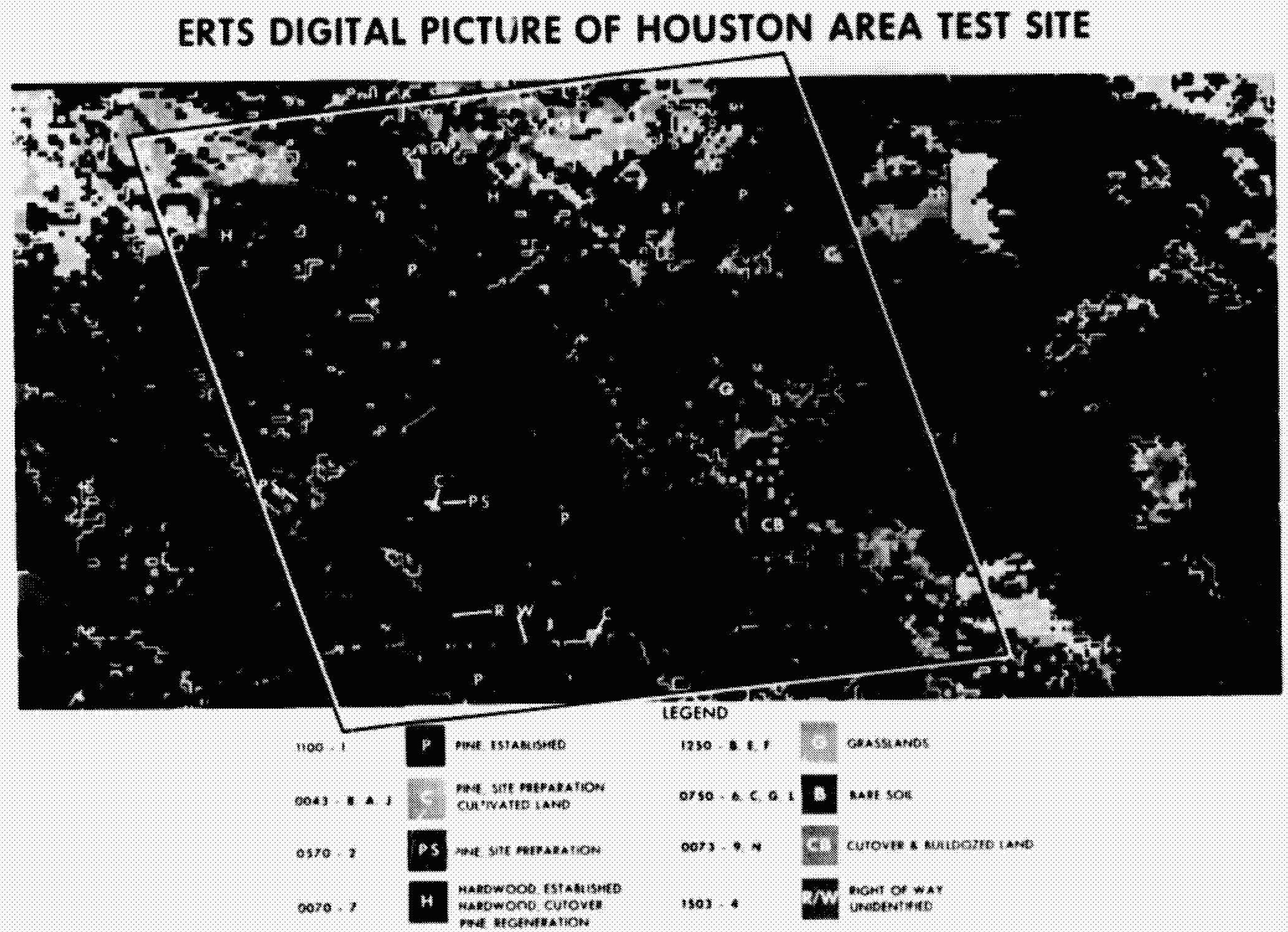


Figure 6-8.- DAS recording of cluster classification map of primary area.

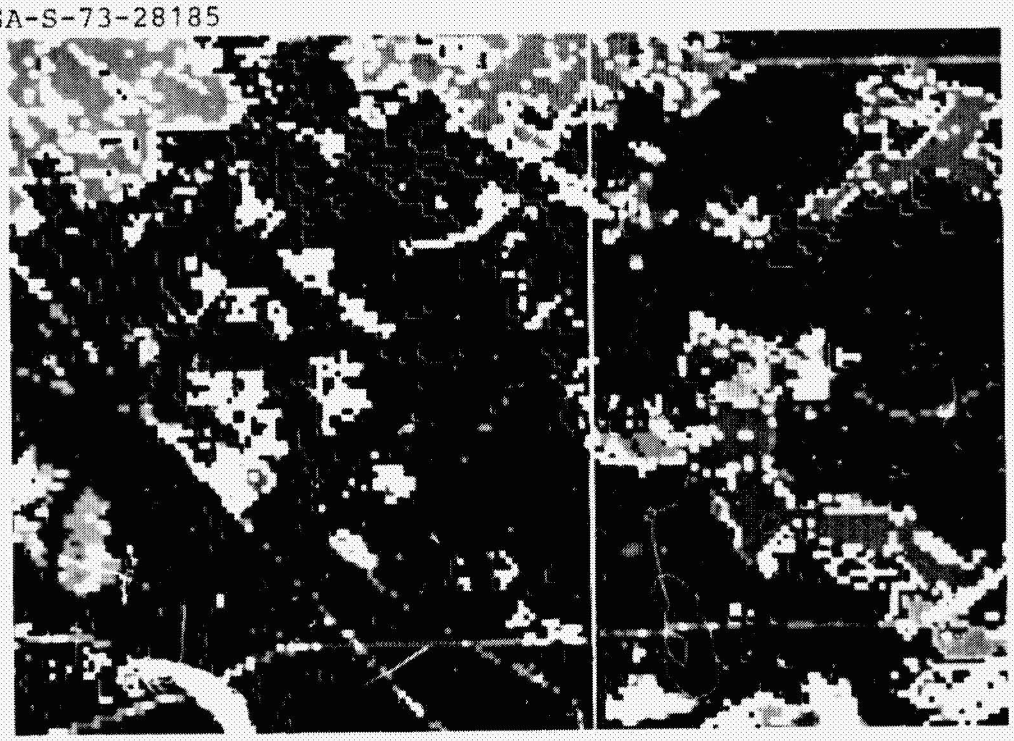
This method of classification was a significant improvement because the reproduction of classes on a DAS recording was possible, with complete fidelity. Two additional classes of rights-of-way were implicit in the data because of the linear patterns they produced, totaling 11 classes. This recording is shown in figure 6-9 in its uncorrected distorted form. This scheme classified the concrete bridge and trestle across Lake Conroe as a separate class and delineated water bodies extremely well. The work was performed so late in the investigation that a map could not be made from the printout; however, the class areas were computed by counting pixels. The accuracy computed by the area comparison method was 74 percent, using the pixel count. This equaled the large-area method of accuracy but because of the other advantages enumerated, has a decided edge over the large-area method. The statistics of this computer run, shown in appendix tables A-V and A-VI, were particularly good. The distances between clusters, set for a DLMIN of 3.2, were very wide, reaching as much as 20.0 points in some cases, with very few cases as low as 3.2.

The success of this clustering method was considered to be due to the small area of data examined by the computer. The data were directly pertinent to the problem and contained none of the peripheral anomalies which are usually found when extensive areas are examined.

6.3.2 Maximum-Likelihood Classification Results

Computer classification results.— The LARSYS program was first submitted using five classes: (1) pine, (2) hardwood, (3) cutover timber, (4) grass, and (5) bare soil. Although the computer line-printer classified map looked good, the

NASA-S-73-28185



LEGEND

- Pine, established
- Hardwood, established
- Grass
- Pine, site prepared/vegetated
- Pine, cutover
- Bare soil
- Pine, regenerated
- Bridge trestle
- Water

Figure 6-9.- DAS recording of a small-area cluster classification map of the primary area.

training field statistics were low and the overall performance was only 183 percent.[4] The team thought that better results could be obtained if more homogeneous training fields were selected with a larger number of classes. The results of the run with improved training fields and more classes were excellent.

The training fields were modified and new classes were added for a total of 12 classes: (1) pine, established; (2) pine, cutover; (3) pine, site prepared/vegetated; (4) pine, regenerated; (5) pine, site prepared; (6) hardwood, established; (7) hardwood, cutover; (8) grass; (9) weeds; (10) cut and bulldozed; (11) water; and (12) bare soil.

An overall performance of 95 percent was achieved with these 12 classes. This performance was excellent; however, the classifications bare soil, pine, site prepared, grass, and weeds were poor. The analysis of the classified data showed that bare soil and pine, site prepared areas, and the grass and weeds areas could not be separated with any consistency. The final classification was made using only 10 classes. The bare soil and pine, site prepared, were grouped into one class called bare soil; the grass and weeds areas were grouped into one class, grass. In addition, two classes of rights-of-way were defined in the printout by their linear pattern. The 10-class computer run improved the performance statistics, showing an overall rating of

---

[4]Performance, when dealing with LARSYS, refers to the accuracy of classification within the training fields and not to the map as a whole. It is a measure of homogeneity of individual pixel values within a field, which by accuracy standards may be already 100-percent correct.

96.1 percent. These statistics are shown in detail in table A-VII in the appendix.

A test classification run was then made using the above 10 classes and training fields, including 18 test fields which contained the 10 categories. The overall performance achieved by the test fields was 95.7 percent, as shown in table A-VIII. A certain amount of operator bias probably occurred by selecting test fields in areas of high homogeneity, which resulted in this extremely high performance within the test fields. This is not related to classification accuracy, which is concerned principally with the definition of boundaries between classes.

When the computer printout was converted to a cartographic format to match the ground-truth map, and the class boundaries were smoothed to make a more usable product, the result was as shown in figure 6-10.[5] An accuracy of 70 percent was achieved when compared with ground truth.

The DAS film output of this computer run made an excellent presentation, except that class boundaries were ragged where the map classes were smoothed, and skew and unequal horizontal and vertical scale distortions were present. A correction for these distortions was in progress but could not be completed in the time allotted. The distorted film output map is shown in figure 6-11.

---

[5]It is felt that there was no loss in accuracy in smoothing because an average line of demarcation was selected, with some pixels included and some excluded. This hypothesis was not tested, however.

NASA S-73-28190

LARSYS CLASSIFICATION MAP
PRIMARY AREA

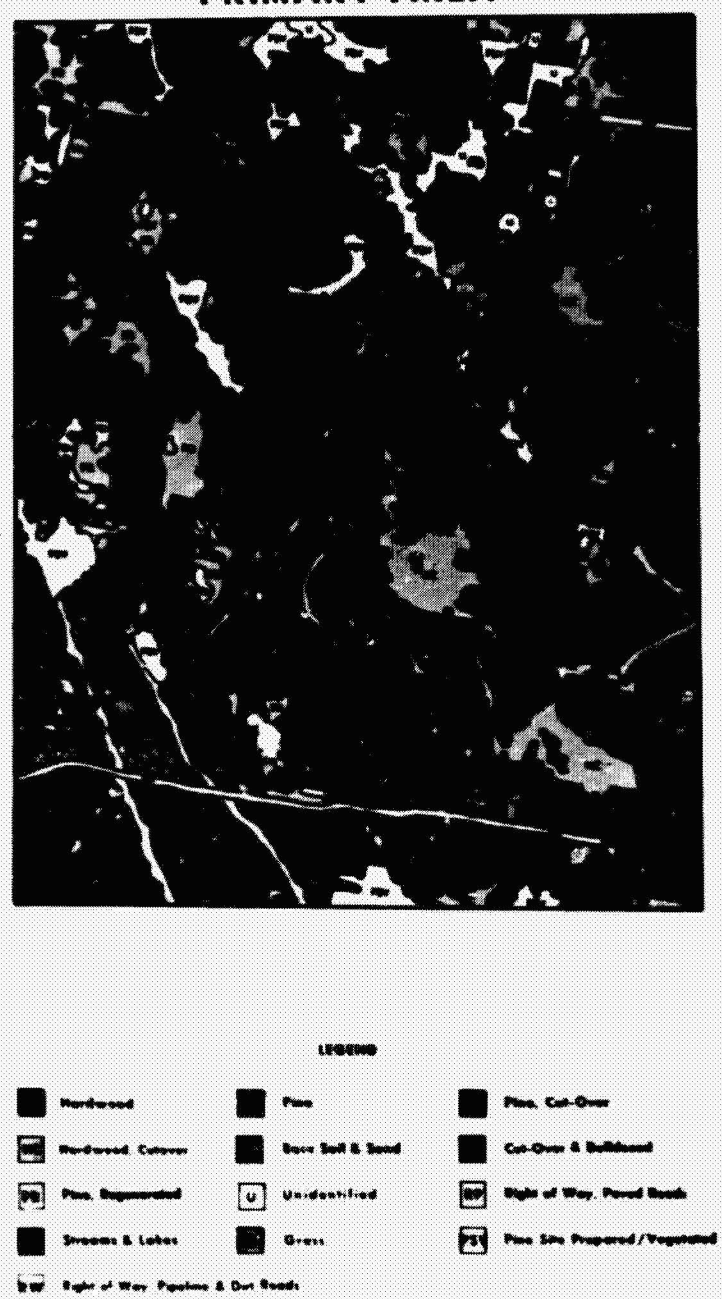
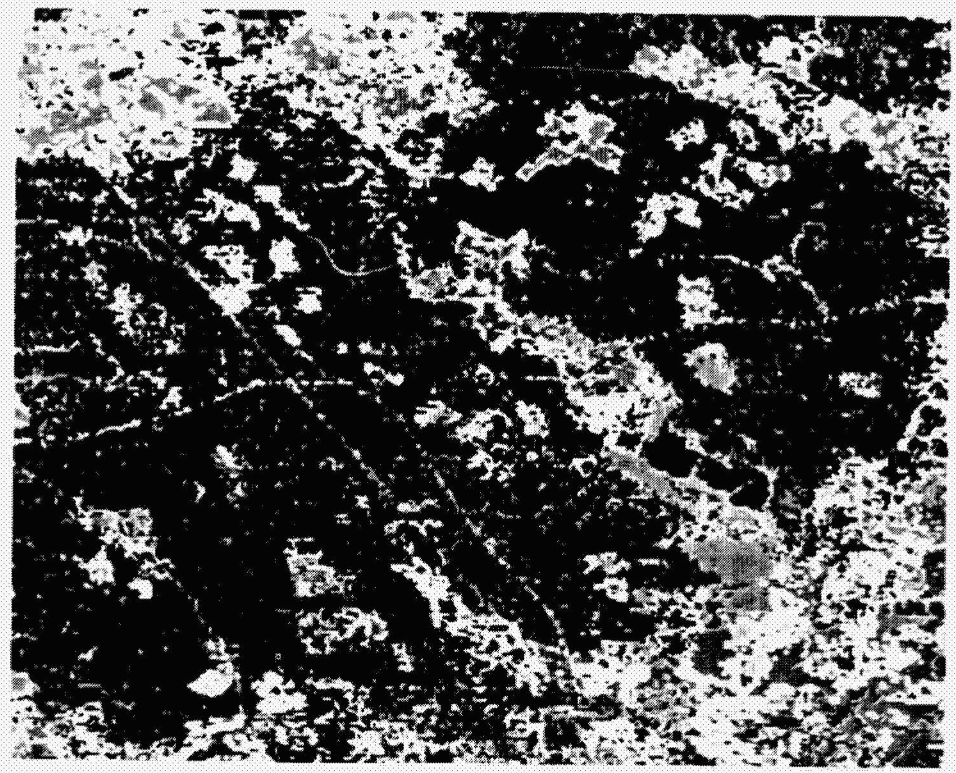


Figure 6-10. – LARSYS classification map of primary area. (Class lines in this map were smoothed and do not represent the actual locations where pixel changes occurred.)

NASA-S-73-28184



LEGEND

Pine, established	Hardwood, cutover
Pine, site prep/veg	Cut and bulldozed
Pine, regenerated	Grass
Pine, cutover	Bare soil
Hardwood, established	Water

Figure 6-11.- DAS recording of LARSYS classification map of primary area.

ERIPS classification results.– At this stage the ERIPS system has not proven its usefulness in forest classification using ERTS-1 data. One of the theoretical advantages is its speed in selecting and delineating training fields for computer classification. This has been proven, but accuracy is lacking in selecting good forest training fields, as evidenced by the poor statistics. To determine if this was caused by software problems, training field coordinates from a LARSYS 1108 printout were entered into the ERIPS; the resulting statistics were the same as those computed by the 1108, which indicated that the method of training field selection was at fault and not the program. The inability to delineate the field boundaries accurately was apparently due to the use of only one channel for gray-mapping, and to the natural character of forest features; i.e., poorly defined, irregular, oddly shaped boundaries, and the erratic operation of the Grafecon pen. Many of these class boundaries are very difficult to define and differences stand out only when multispectral analysis is complete. They do not stand out well in any single channel. This is inherent in a multispectral system, otherwise classes could be outlined using only one channel. The problem is maximized with forest classes because of their gradual gradations and vague boundaries, some of which are visible on one channel and some on another. The best run produced 12 classes, with an overall performance rating of only 61.3 percent.

6.4 SIGNATURE EXTENSION RESULTS

6.4.1 Conventional Classification Results

Single-band classification.– The secondary area was classified using the same density signatures found in the

primary area to delineate classes. The accuracy of the classifications shown in figure 6-12 was also checked by superimposing the two maps and counting the points of agreement versus disagreement and by area comparison. The mean accuracy was 73 percent. The ground-truth map against which these data were compared is shown in figure 6-13.

Optical multiband classification results.— Optical multiband classification of the secondary area was done in the same way as the single band classification. The results were 74-percent mean accuracy.

MCFV classification results.— The secondary area was also classified in the same way using the results of analysis on the multichannel film viewer. The results were 58-percent mean accuracy.

JSC color composite classification results.— The secondary area was also classified in the same way using the results of the analysis in the primary area of the data analysis station (DAS). The results were 78-percent accuracy.

6.4.2 Computer Classification Results

Cluster classification results.— The extension of the signatures of a clustering program appears to be generally valid for areas in the secondary area as long as both areas are on the same data set. However, if the same parameters are used to define and limit the analysis of data on an adjacent data set, different clusters appear because the computer has examined different data. This situation usually occurs when additional lines or columns are included

NASA S-73-25507

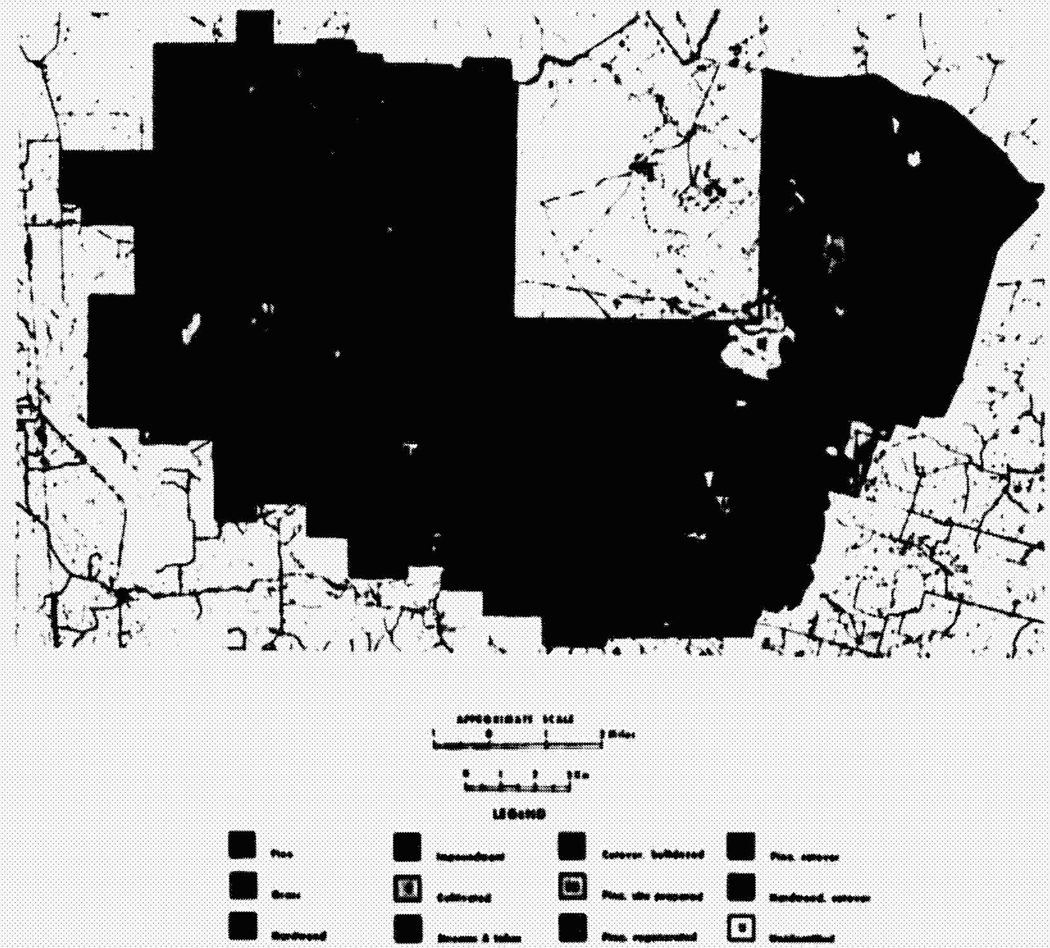


Figure 6-12.- Single-band conventional classification map of secondary area.

NASA S-73-28186

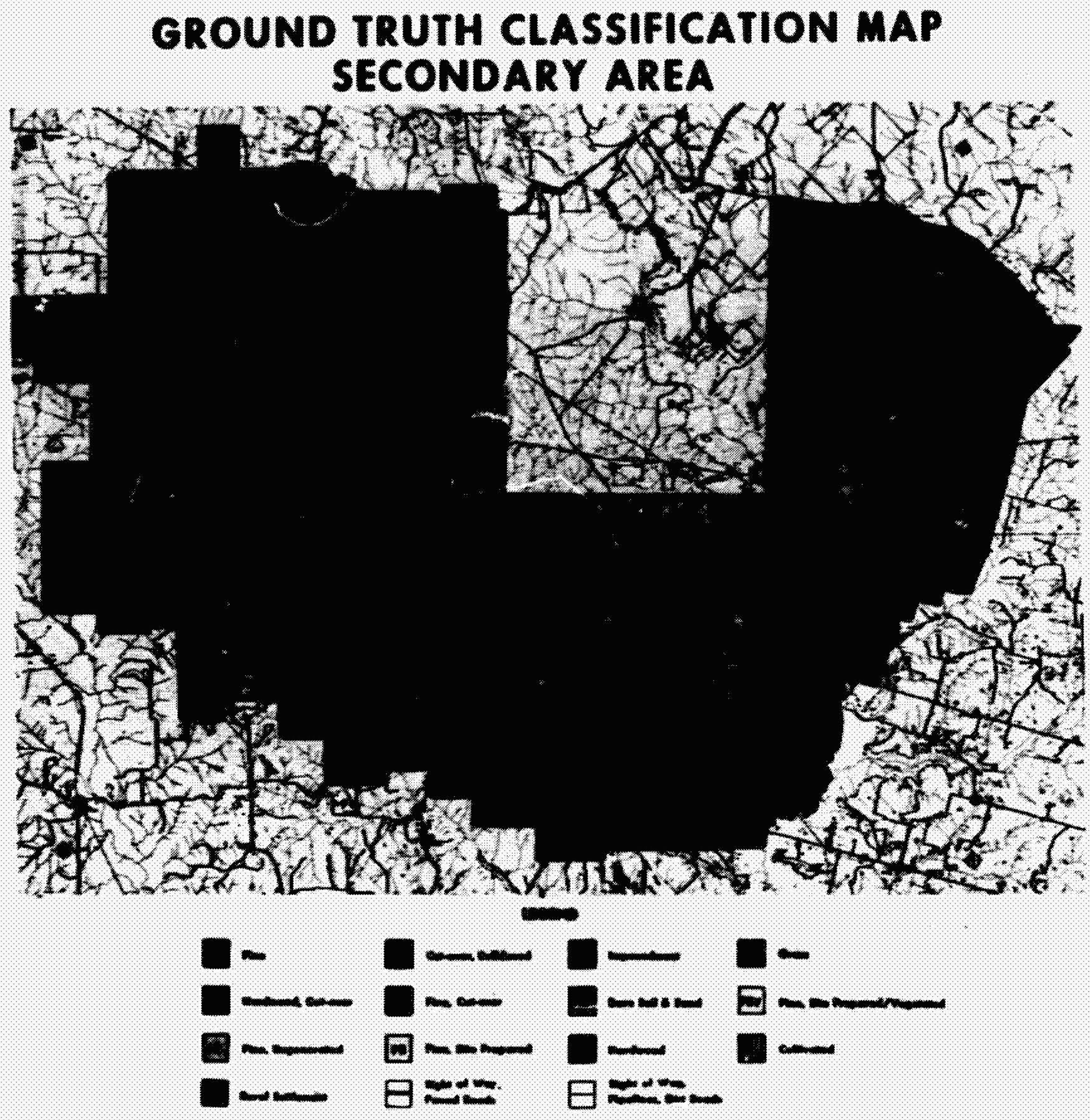


Figure 6-13.- Ground-truth classification map of secondary area.

on a tape or when an adjacent tape is analyzed. When the subject for analysis falls on two data tapes, the tapes are recommended to be merged so that the data may be processed as one data set. This limitation impeded the extension of information using a clustering program. When the data were confined to one tape the results were 68-percent accurate, as shown in figure 6-14.

Maximum-likelihood classification results.— The team was not able to complete a LARSYS classification map of the secondary area because of lack of time. However, the data were available for study in printout form. The data were generally good in classes similar to those in the primary area, but degraded in peripheral areas where classes were somewhat different and lacked training fields.

The performance of LARSYS training fields in the primary area was tested by selecting test fields in the secondary area. This constituted an indirect test of signature extension, and the results showed 95.7-percent accuracy. (Paragraph 6.3.2 provides a discussion of these results.)

6.4.3 Information Extension Results

The accuracies cited above were generally comparable, although somewhat higher than those for the primary study area. (Table 6-V provides a comparison of results.) This proves that signature extension under these conditions is valid. Two possible explanations exist for the higher accuracy rating which resulted in the secondary area. One is that the smaller scale maps used for classification and

NASA S-73-28196

CLUSTER CLASSIFICATION MAP SECONDARY AREA

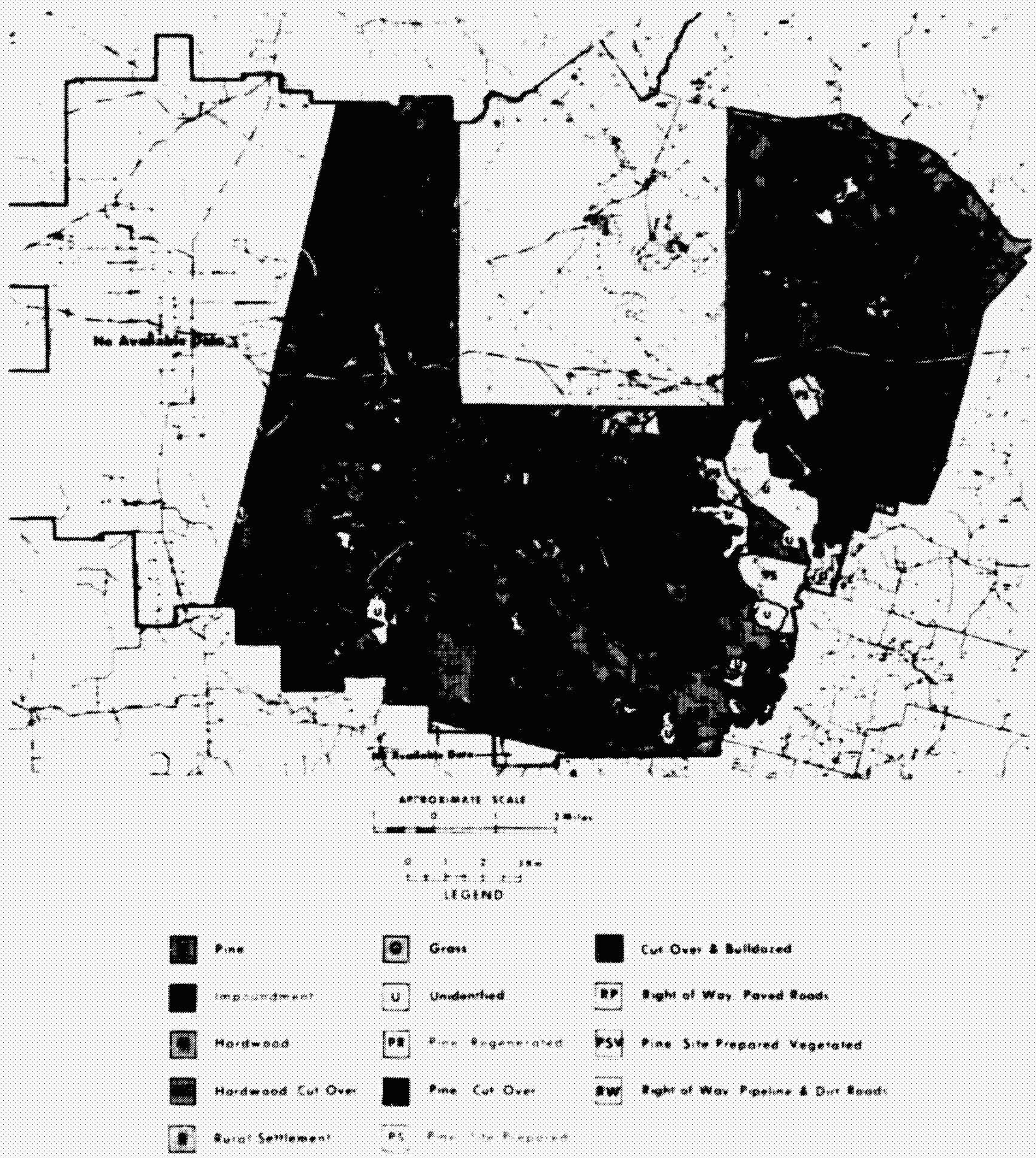


Figure 6-14.- Classification map of secondary area derived from clustering.

TABLE 6-V.- SUMMARY OF ACCURACIES OF CLASSIFICATIONS USED

Primary Area	Accuracy by Grid Count (percent)	Accuracy by Area Computation[a] (percent)	Mean Accuracy	No. of Classes	Classification (percent)	Achievement (percent)
Single Band	66	58	62	7	50	56
Optical Multiband	55	64	60	10	72	66
I^2S Multiband	56	82	69	7	50	60
DAS Multiband	--	67	67	10	72	70
Clustering (large area)	66	81	74	14	100	87
Clustering (small area)	--	74	74	11	79	77
LARSYS	67	73	70	12	86	78
Secondary Area						
Single Band	63	83	73	11	79	76
Optical Multiband	63	84	74	10	72	73
I^2S Multiband	56	60	58	7	50	54
DAS Multiband	--	78	78	10	72	75
Clustering	60	75	68	14	100	84

[a]Explanation of accuracy measurement approach given in paragraph 5.5.

for ground truth resulted in a smaller sample. Secondly, the two classification maps for single-band and optical-multiband analysis were examples of the error possible when only area accuracy is determined.

6.5 ACCURACY EVALUATION RESULTS

Table 6-V summarizes the accuracies of the investigation performed by the analysis team. (Refer to paragraph 5.5 for a discussion of the accuracy measurement approach.)

6.5.1 Accuracy Evaluation Results

Because accuracy figures are not entirely reflective of achievement, an achievement rating was computed for each method. This was necessary because the mean accuracy figures contained the lack of precision inherent in the area comparison method of computation. Secondly, mean accuracy failed to reflect the number of classes achieved. The achievement rating therefore was computed by averaging the mean accuracy percent and the percent of possible classifications achieved. If only seven of a possible 14 classes are achieved, the classification accuracy is 50 percent.

$$\text{Achievement} = \frac{\text{Accuracy \%} + \text{Classification \%}}{2}$$

This favors the method which achieves a large number of classes while still maintaining a high level of accuracy.

The accuracy figures do not reflect what could have been achieved if only the second level of the hierarchy or some other simplified scheme had been attempted in classifying

forest features. An examination of the data suggests that much higher accuracies would have been possible. Even with the complications of using 14 classes, pine, established, was classified with 91-percent accuracy when only the computer methods were considered, as shown in table 6-VI.

TABLE 6-VI.- ACCURACY OF INDIVIDUAL CLASSES

Class	Average Accuracy by Conventional Methods	Average Accuracy by Computer Methods
	%	%
Pine, established	93	91
Grass	60	85
Hardwood, established	61	70
Pine, regenerated	31	31
Pine, cutover	13	41
Pine, site prep./veg.	0	0
Hardwood, cutover	42	56
Bare Soil	0	0
Cut and bulldozed	16	0
Rural settlement	0	22
Lakes	0	0
Rights-of-way/utility lines	0	36
Rights-of-way/paved	10	47

Grass was classified with 85-percent accuracy, and hardwood, established, with 70-percent accuracy. These figures were computed by determining the mean accuracy of each of these classes separately from table A-I for conventional and for computerized processing methods. The accuracies of other classes were generally much lower by both conventional and computer methods. If this limited number of classes had been processed singly, the accuracies would probably have been much higher. This capability suggests that computer analysis could be used for differentiation of small numbers of classes, such as the first and second levels, for forest surveys with a mean accuracy of at least 80 percent.

A study of the accuracy figures of all the classification methods (shown in tables A-I and A-II in appendix A) indicates that the classes which contributed most to the classification errors were pine, regenerated; pine, site prepared/vegetated; pine, cutover; and hardwood, cutover. These classes have complex spectral signatures which can easily be confused with other classes. Bare soil, cut and bulldozed, and pine, site prepared, were also frequently confused with each other. Rural settlements were difficult to classify due to the small spatial coverage of the buildings which delineates this class from grass. Since cultivated land occurred only in the secondary area, no firm evaluation of this class was possible.

7.0 ADDITIONAL EXPERIMENTAL WORK

As the investigation ended, the team was studying several advanced computer techniques to permit registration of the data and to provide temporal studies.

One of the registration techniques shows considerable promise, that of registering data from two time periods and using the eight channels to produce a classification map. Using the same training fields previously used for four channels, the classification performance was increased for all but one class.

The important task of registering the DAS film output to a map was not undertaken because the film transport lacked constant speed. This needs correction.

Another advanced technique experimented with was the removal of stripes from the data. Some of the 24 detectors (four channels with six detectors each) in the multispectral scanner have not functioned properly and they register a bright striping effect in the imagery. This can be seen in all of the ERTS-1 multispectral scanner data, even those that have been processed through supervised and unsupervised classification programs. The striping effect is a hindrance to interpretation and conventional and digital analysis, and reduces classification accuracy to some extent. Analysts at JSC have worked out a smoothing program which averages the spectral data of all the detectors over the entire scene. This is done without changing the radiance values, thus preserving the fidelity of the data. The team was able to

make use of the technique just before the final report, but unfortunately there was not time to follow it through to completion. Smoothing is anticipated to improve the performance of the clustering and maximum-likelihood programs, and this should be extended in subsequent investigations.

8.0 DISCUSSION OF TECHNIQUE DEVELOPMENT

The unique approaches and techniques developed for interpretation of the ERTS-1 data greatly influenced the accuracy shown on the previous pages. The following factors have a bearing on how these figures should be interpreted.

1. The procedures used in this investigation were largely new and experimental. In each case much better results and levels of accuracy could have been achieved if additional adjustments had been made. Time and manpower for such improvements were not available.

2. A comparison of the results of these procedures with one another was difficult because the parameters were not held constant throughout the study. The number and kinds of classes changed from time to time and different techniques were experimented with.

3. The team spent very little time in the field for ground-truth. This seemed inadequate for securing an accurate picture of ground conditions. The ground-truth map was the standard against which accuracies were determined and should have been developed to a high degree of accuracy. However, the accuracy of this ground-truth map is not known. As a result, part of the classification inaccuracy may have been due to comparing correct classifications with incorrect ground truth. A statistical sampling of the ground-truth map should have been conducted to determine its accuracy, although this would have been a rather large and expensive field operation.

4. No precedent was known for comparing classification maps with ground truth and determining their accuracy. As a result, the two methods used were developed experimentally. The inconsistency of the results suggests there was inadequacies in one or both of these methods.

5. The forest investigated did not lend itself well to exact classification. The timber classes existed in mixtures of species, age, vigor, site, and size classes which blended gradually from one to another. It was extremely difficult to draw boundary lines which would remain constant from one classification technique to another.

6. The significance of achieving only 70- to 74-percent accuracy in computerized processing should not be minimized. The matching of 14 very complex classes is not easily achieved by any method, and this range of accuracies is therefore regarded as high. Generally, if fewer classes had been used a much higher accuracy could have been achieved.

7. This investigation focused on whether ERTS-1 data were suitable for intensive forest surveys. The study of intensive forest surveys had been required by the number and type of forest features originally specified. These are the kinds of surveys of most interest to American foresters, because American forestry has developed to the point that intensive management is now necessary. However, there is still a very important role for extensive surveys

which encompass entire working circles, counties, states, and regions. Undeveloped countries have a far greater interest in this type of information than in the details of intensive surveys.

9.0 CONCLUSIONS AND RECOMMENDATIONS

9.1 CONCLUSIONS

1. The team met the four main objectives of the investigation:

    a. detection evaluation,

    b. classification evaluation,

    c. information extension evaluation, and

    d. accuracy evaluation.

2. The ERTS-1 data can probably be used best in forestry applications if applied to extensive surveys where a few broad generalized classifications are needed, rather than for intensive surveys where many detailed stand condition classes are required. When a large number of classes were attempted, the accuracies achieved did not consistently reach the criterion of 80-percent accuracy or better. However, this study indicates that a few major features may be classified using ERTS-1 data with a mean accuracy greater than 80 percent. ERTS-1 data could quite possibly be used as one component in a multistage sampling technique to greatly simplify intensive aircraft surveys.

3. The classes from the original hierarchy which were mapped were:

Pine, established	Cultivated
Hardwood, established	Grass
Pine, cutover	Rights-of-way, roads and utility lines
Pine, site prepared	Rights-of-way, paved roads
Pine, site prepared/ vegetated	Cut and bulldozed
Pine, regenerated	Bare soil
Hardwood, cutover	Rural settlement
Lakes	

4. The clustering and maximum-likelihood methods of classification were both efficient. The differences in accuracies achieved were not significant enough to specify one over the other, because of the imprecise methods of accuracy measurement. However, computer methods as a group can probably be rated superior to conventional methods for forest classification. The achievement rating, which considers the number of classes as well as the accuracy, is probably a better index of efficiency than accuracy only.

5. In addition to the possibilities for producing the timber inventory maps described above, the ERTS-1 data have the additional advantage of providing sequential coverage, which is provided at 18-day intervals. One coverage can be registered to another either visually or by computer by the identical scale and time of coverage. Registering provides the opportunity of doing continuous temporal studies of changing situations, such as the progress of sites being prepared for planting, timber cutting, and detecting insect and disease epidemics, which has never before been possible.

9.2 RECOMMENDATIONS

1. Further study is recommended in all phases of this investigation. More work is needed, particularly in evaluating

   a. seasonal and atmospheric effects,

   b. registration corrections, especially imagery to map coordinates by means of the DAS,

   c. temporal correlations,

d. JSC color composite classification of forest features, and

e. the removal of striping effects.

2. Greater emphasis should be placed on securing accurate ground-truth maps in future investigations and in determining their accuracy.

3. A better method of statistically evaluating the accuracy of classification maps should be developed.

4. A study should be made to determine how smoothing of classification lines affects the accuracy of thematic maps.

Lyndon B. Johnson Space Center
National Aeronautics and Space Administration
Houston, Texas, January 21, 1974
641-14-07-50-72

10.0 REFERENCES

1. Adapted from a Forest Service handbook on forest stand and condition classes for the Southern Region, which were in turn adapted from the standard reference work by the Society of American Foresters, Forest Cover Types of North America, Washington, DC, 1954.

2. Aldrich, R. C.: Classifying Forest and Nonforest Land on Space Photographs. 3rd Annual Earth Resources Program Review, JSC, Houston, TX, Dec. 1-3, 1970.

3. Heyning, J. M.: The Human Observer. Seminar Proceedings The Human in the Photo-optical System, Soc. of Photo-Optical Instrumentation Engineers, New York, NY, Apr. 25-26, 1966.

4. Kan, E. P.; and Holley, W. A.: ISOCLS (ISODATA) Clustering, a Well-defined Problem. LEC/HASD 640-TR-152, Lockheed Electronics Co., Inc., Houston, TX, Dec. 1972.

5. Hoffer, R. M.; Anuta, P. E.; and Phillips, T. L.: ADP, Multiband, and Multiemulsion Digitized Photos. Photogr. Engr., vol. XXXVIII, no. 10, Oct. 1972, pp. 989-1001.

APPENDIX A

OTHER SIGNIFICANT FINDINGS

During the course of the investigation a number of incidental items of interest to foresters were found in the data. As a result, the team published two significant findings reports. The first reported the detection of small ponds, the details of which are covered in paragraph 6.1, Evaluation of Detectable Area Size. The second report covered detection of the effects of a forest fire in the national forest which covered approximately 40 hectares (100 acres). This was a light ground fire, which was a prescribed burn to clear brush before marking timber for cutting. The fire occurred in mid-October 1972, and was detected on the next coverage, November 27, 1972. By this time many of the pine needles on the trees had turned brown. These effects caused the area to appear as a black smudge in otherwise red coloration (which indicates healthy trees) on EBR type-C composites. A JSC enhancement of this area is shown in figure A-1, which shows the burn as a reddish brown area. The precision with which the effects of this light ground fire were registered indicates that ERTS data may be used in the future in fire damage assessment, which requires the mapping of the perimeter of large fires.

During the investigation a pine-bark beetle epidemic occurred within the test area, and for this to have been detected would have been a finding of considerable significance. However, the trees killed occurred in small groups,

NASA S-73-25467

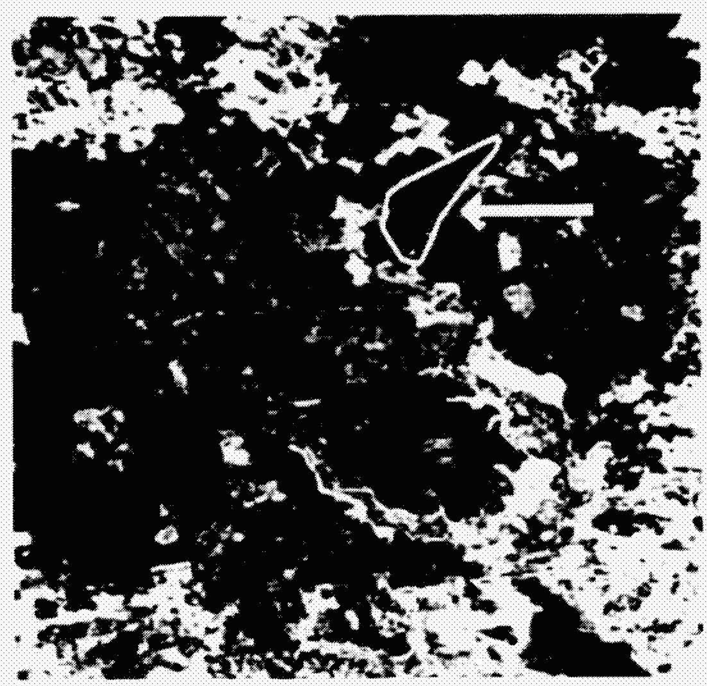


Figure A-1.- JSC enhancement of light ground fire in the Sam Houston National Forest.

usually less than a hectare in size, and these groups were smaller than the ERTS-1 resolution limit. The damage was detected indirectly in one case, however, where approximately 2 hectares of pine were killed in a stand which was classified as hardwood, although nearly 50 percent of the stand was actually pine. The LARSYS program classified this portion of the stand as "hardwood, cutover." Because this is a supervised program a beetle-infested training field would have to have been included for a direct classification to have been possible. None of the other classification systems indicate this phenomenon.

TABLE A-I.- COMPUTATIONS OF ACCURACY BY THE AREA DIFFERENCE METHOD IN THE PRIMARY STUDY AREA

Classification	Ground Truth (hectares)	Single Band		Optical Multiband		MCFV		JSC Composite		Cluster (L.A.)		Cluster (S.A.)		LARSYS	
		Area	Diff	Area	Diff	Area	Diff	Area	Diff	Area	Diff	Area	Diff	Area	Diff
Pine, established	3,338	3,587	249	3,299	39	3,517	179	3,785	447	3,430	92	3,845	507	3,042	298
Pine, regenerated	378	287	91	74	304	636	256	-	378	27	351	541	163	245	133
Pine, cutover	22	-	22	9	13	-	22	[a]107	415	1	21	[a]161	355	16	6
Pine, site prep/veg	103	17	86	376	273	-	103	-	103	158	55	150	47	313	210
Hardwood	826	1,700	874	1,021	186	596	230	818	8	637	189	723	103	1,288	461
Hardwood, cutover	502	-	502	156	346	522	20	-[a]	-	746	244	-[a]	-	308	194
Grass	815	413	402	457	358	753	62	1,309	493	757	58	663	152	649	166
Bare soil	25	-	25	581	556	-	25	19	6	31	6	96	71	95	70
Cut and bulldozed	38	-	36	-	38	-	38	21	17	169	131	-	38	41	3
Cultivated	-	21	21	-	-	-	-	-	-	-	-	-	-	-	-
Rural settlement	14	-	14	-	14	-	14	-	14	21	7	-	14	-	14
Impoundments	10	2	8	-	10	-	10	77	67	20	10	-	10	-	10
Bridge	-	-		-	-	-	-	-	-	-	-	6	6	-	-
Lakes and streams	3	-	3	-	3	-	3	-	3	-	-	5	2	4	1
Rights-of-way/util	44	-	44	70	26	134	90	62	18	51	7	-	44	78	34
Rights-of-way/paved	79	-	79	165	86	39	40	-	79	54	25	-	79	56	23
Unidentified	-	172	172												
Total	6,198		263		2,252		1,094		2,048		1,198		1,591		1,686
Percent error			42%		36%		18%		33%		19%		26%		27%
Percent accuracy			58%		64%		82%		67%		81%		74%		73%

[a]Pine, cutover, and Hardwood, cutover, were combined.

TABLE A-II.- COMPUTATIONS OF ACCURACY BY THE AREA DIFFERENCE METHOD IN THE SECONDARY STUDY AREA

Classification	Single Band			Optical Multiband			MCFV			JSC Color Composite			Cluster		
	Ground Truth (Hectares)	Area	Diff	Ground Truth (Hectares)	Area	Diff	Ground Truth (Hectares)	Area	Diff	Ground Truth (Hectares)	Area	Diff	Ground Truth (Hectares)	Area	Diff
Pine established	21,445	21,933	488	20,476	20,151	325	21,006	18,321	2,685	15,209	16,206	997	15,209	15,065	144
Pine regenerated	447	227	220	252	57	195	422	2,760	2,338	384	-	394	384	150	234
Pine, cutover	911	555	356	887	1,029	142	858	-	858	846	297	549	846	1,064	218
Pine, site prepared	339	-	339	322	277	45	326	-	326	246	-	246	246	332	86
Pine, site prep/veg	355	-	355	355	-	355	355	-	355	254	-	254	254	170	84
Hardwood	2,998	2,687	311	2,852	3,051	199	2,873	2,516	357	2,338	1,988	350	2,338	931	1,407
Hardwood cutover	325	939	614	301	774	473	272	3,096	2,824	309	297	12	309	1,493	1,184
Grass	2,064	2,930	866	1,619	1,238	381	1,858	2,465	607	1,429	2,685	1,256	1,429	1,522	93
Bare soil	22	-	22	22	1,479	1,457	22	-	22	22	170	148	22	-	22
Cut and bulldozed	989	231	758	989	21	968	989	-	989	989	412	577	989	251	739
Cultivated	42	113	71	2	-	2	42	-	42	42	-	42	42	-	42
Rural settlement	24	-	24	24	-	24	24	-	24	24	-	24	24	99	75
Impoundment	125	4	121	73	97	24	109	-	109	89	125	36	36	12	77
Unidentified	-	467	467											1,093	1,093
Total	30,086		5,012	28,174		4,590	29,156		11,536	22,181		4,875	33,181		5,497
Percent error			17%			16%			40%			22%			25%
Percent accuracy			83%			84%			60%			78%			75%

TABLE A-III.- STANDARD DEVIATIONS OF THE CLUSTERS AND THE DISTANCES BETWEEN CLUSTER MEANS OF THE FIRST NINE CLUSTERS OF THE LARGE-AREA CLUSTERING CLASSIFICATION

STANDARD DEVIATIONS					
Cluster	Class	Channel 1	Channel 2	Channel 3	Channel 4
1	Pine, established	1.12	.77	1.95	1.20
2	Pine, site prepared	1.20	1.32	1.42	1.17
3	Cut and bulldozed	1.24	1.81	1.99	1.26
4	Rights-of-way	1.12	1.77	1.32	.90
5	Unidentified	1.39	1.96	1.71	1.05
6	Impoundments	1.60	2.51	1.44	.99
7	Hardwood, established	1.06	.99	1.79	1.15
8	Cut and bulldozed	1.32	1.47	1.28	.98
9	Unidentified	1.83	1.98	1.95	.89
A (10)	Bare soil	1.29	1.54	1.97	1.39
B (11)	Grass	.84	1.35	1.46	1.19
C (12)	Unidentified	1.59	2.38	1.63	1.12
D (13)	Impoundments	2.40	3.12	1.97	1.29
E (14)	Bare soil	1.07	1.42	1.51	1.17
F (15)	Grass	1.06	1.79	1.48	1.20
G (16)	Bare soil	1.46	1.74	1.83	1.06
H (17)	Impoundments	1.71	2.84	2.75	1.64
I (18)	Unidentified	1.31	1.66	1.36	1.10
J (19)	Rural settlement	1.32	1.23	1.36	1.04
K (20)	Bare soil	1.24	1.88	2.37	1.22
L (21)	Unidentified	3.91	5.35	3.95	1.62
M (22)	Rights-of-way/paved	1.50	2.63	1.44	1.00
N (23)	Unidentified	1.42	1.47	1.86	.97
O (24)	Unidentified	1.21	1.63	1.36	.83

CLUSTER MEANS				
Cluster	CH(1)	CH(2)	CH(3)	CH(4)
1	24.00	13.92	33.68	18.93
2	27.26	19.37	41.44	22.96
3	32.40	25.01	52.20	28.32
4	26.14	17.95	37.34	20.64
5	33.47	29.53	46.57	24.45
6	39.11	37.67	56.00	28.25
7	24.51	14.53	38.86	22.27
8	27.83	19.18	45.37	25.46
9	44.02	47.18	52.01	23.25
10	29.57	19.94	53.73	30.16

TABLE A-IV.- PARTIAL RESULTS OF DISTANCES BETWEEN CLUSTERS IN LARGE-AREA CLUSTERING CLASSIFICATION

Cluster	1	2	3	4	5	6	7	8	9	10
1	.00	8.40	16.91	4.85	17.12	25.89	4.06	11.20	32.40	15.19
2	8.40	.00	9.58	3.98	8.70	17.39	5.17	3.27	21.69	9.45
3	16.91	9.58	.00	13.41	5.19	7.93	13.55	7.10	14.32	4.05
4	4.85	3.98	13.41	.00	11.23	20.71	3.53	8.18	22.59	13.59
5	17.12	8.70	5.19	11.23	.00	8.80	13.90	7.49	11.56	8.71
6	25.89	17.39	7.93	20.71	8.80	.00	22.04	14.90	8.51	11.26
7	4.06	5.17	13.55	3.53	13.90	22.04	.00	7.08	28.[illegible]	11.74
8	11.20	3.77	7.10	8.18	7.49	14.90	7.08	.00	19.96	6.76
9	32.40	21.69	14.32	22.59	11.56	6.51	28.24	19.90	.00	18.61
10	15.19	9.45	4.05	13.59	8.71	11.26	11.74	6.76	18.61	.00
11	12.49	3.85	7.25	7.01	5.85	15.10	9.49	4.06	19.64	8.38
12	22.55	13.57	8.67	15.10	4.74	7.30	19.26	12.40	7.02	12.67
13	26.69	18.42	10.50	20.53	10.12	3.24	23.12	16.38	3.54	13.62
14	14.74	6.46	3.74	10.45	4.74	11.71	11.19	4.03	17.45	5.33
15	14.47	6.38	7.23	8.44	3.48	12.85	11.73	6.59	15.16	9.74
16	21.02	12.25	4.57	15.16	3.77	5.06	17.47	10.25	10.05	8.77
17	11.57	7.96	10.78	7.11	7.84	14.78	10.87	10.16	14.39	12.53
18	29.18	18.63	13.57	19.03	8.92	10.38	25.50	17.73	6.84	18.12
19	14.08	7.08	5.10	11.62	8.14	13.51	9.98	3.64	20.52	3.57
20	18.59	10.90	9.79	11.17	5.34	11.87	16.32	11.43	12.07	13.13
21	29.43	21.87	14.70	23.07	14.58	8.74	26.10	20.13	7.87	17.06
22	9.30	5.25	11.11	4.84	7.36	16.52	8.38	8.36	16.04	12.82
23	31.73	20.43	12.93	21.35	9.52	5.75	27.54	18.79	3.52	17.75
24	25.76	16.10	13.40	16.28	8.02	13.13	22.57	16.09	10.26	17.69
25	3.70	5.40	12.30	2.88	11.61	19.15	4.57	8.37	22.63	11.59
26	26.81	17.31	8.82	19.99	7.36	3.20	23.06	15.10	5.96	12.97

TABLE A-V.- STATISTICS SHOWING THE STANDARD DEVIATIONS OF THE CLUSTERS AND DISTANCES BETWEEN CLUSTER MEANS OF THE SMALL-AREA CLUSTERING CLASSIFICATION

Standard Deviations

Cluster	Class	Channel 1	Channel 2	Channel 3	Channel 4
1	Pine, established	3.77	4.71	2.67	1.87
2	Hardwood, established	.98	.74	1.13	.92
3	Grass	1.18	1.30	1.07	1.04
4	Grass	1.30	1.68	1.91	1.11
5	Pine, site prepared	1.17	.83	1.52	.94
6	Rights-of-way/paved	.99	.76	1.31	.97
7	Grass	1.10	1.11	1.10	.83
8	Hardwood, cutover	.69	1.23	1.16	.87
9	Grass	1.01	1.64	1.27	.98
A (10)	Bare soil	1.13	1.46	1.39	1.05
B (11)	Rights-of-way/paved	1.15	1.87	2.30	.95
C (12)	Hardwood, cutover	1.27	1.17	1.22	1.04
D (13)	Grass	1.97	3.18	2.39	1.23
E (14)	Grass	1.00	.92	1.23	1.00
F (15)	Grass	1.46	1.20	1.80	1.20
G (16)	Grass	.79	1.22	1.40	.98
H (17)	Unidentified	2.34	3.01	2.56	1.51
I (18)	Unidentified	1.18	1.00	1.42	1.19
J (19)	Unidentified	1.11	1.53	.86	.88
K (20)	Unidentified	1.14	1.59	1.36	1.02
L (21)	Unidentified	2.60	2.75	4.83	4.40

Cluster Means

Cluster	Channel 1	Channel 2	Channel 3	Channel 4
1	42.47	42.60	56.27	27.40
2	24.00	13.99	35.70	20.14
3	30.42	23.64	42.55	23.05
4	31.78	24.16	51.58	28.04
5	23.87	13.82	32.75	18.53
6	24.32	14.29	39.07	22.52
7	29.31	20.88	44.20	24.45
8	25.81	17.24	39.24	21.98
9	25.75	16.60	42.66	24.20
10	31.15	23.74	46.87	25.71
11	32.14	29.12	35.74	18.26
12	19.63	46.69	26.39	
13	36.37	32.94	51.16	25.92
14	27.67	19.71	40.75	22.41
15	29.27	19.77	54.77	30.71
16	25.09	18.10	35.95	28.86
17	34.98	31.62	44.90	22.70
18	29.08	20.09	50.45	27.99
19	29.71	23.03	39.67	21.09
20	28.64	21.51	36.63	19.71
21	29.22	21.33	27.78	12.70

TABLE A-VI.- PARTIAL RESULTS OF DISTANCES BETWEEN CLUSTERS IN SMALL-AREA CLUSTERING CLASSIFICATION

Cluster	1	2	3	4	5	6	7	8	9	10
1	.00	22.28	12.93	8.41	21.69	20.24	13.68	18.13	14.86	10.35
2	22.28	.00	13.38	17.53	2.85	3.77	12.85	5.46	7.76	15.64
3	12.93	13.38	.00	7.93	14.14	11.37	3.28	7.85	6.54	4.41
4	8.41	17.53	7.93	.00	18.06	14.51	7.06	13.01	9.73	3.66
5	21.69	2.85	14.14	18.06	.00	6.16	14.13	7.38	9.69	16.36
6	20.24	3.77	11.37	14.51	6.16	.00	9.86	3.60	4.13	12.88
7	13.68	12.85	3.28	7.06	14.13	9.86	.00	7.30	4.82	3.78
8	18.13	5.46	7.85	13.01	7.38	3.60	7.30	.00	3.72	10.54
9	14.86	7.76	6.54	9.73	9.69	4.13	4.82	3.72	.00	7.68
10	10.35	15.64	4.41	3.66	16.36	12.88	3.78	10.54	7.68	.00
11	12.68	15.06	7.52	12.48	14.29	15.24	10.75	11.54	11.85	10.28
12	12.92	13.38	6.10	5.54	14.67	9.78	3.32	8.44	4.98	4.04
13	4.04	18.67	8.47	5.10	18.62	16.58	9.21	14.51	11.58	6.00
14	16.28	9.21	4.70	10.21	10.37	7.41	4.18	3.49	3.94	7.42
15	11.37	18.33	11.62	4.61	18.99	14.57	9.82	14.29	10.65	7.49
16	18.35	4.78	8.91	14.21	5.46	5.52	9.32	3.39	6.65	11.89
17	6.45	15.01	5.10	6.37	14.90	13.81	7.05	10.91	9.21	5.15
18	12.56	16.20	8.49	3.93	17.05	12.58	6.19	11.41	7.89	4.84
19	15.37	10.87	3.71	11.81	11.72	9.73	6.31	6.22	7.04	8.27
20	15.52	8.19	6.26	12.52	8.39	8.42	8.07	5.44	7.59	9.81
21	11.92	7.87	8.22	10.65	6.76	8.71	9.40	7.57	8.72	9.68

ORIGINAL PAGE IS OF POOR QUALITY

TABLE A-VII.- LARSYS TRAINING FIELDS

Class Name	Field ID	No. Points	Misclassified	% Correct	% Area
Pine, established	P1	132	1	99.2	
	P2	72	0	100.0	
	P3	200	5	97.5	
		404	6	98.5	45.37
Pine, site prepared/ vegetated	PSV1	6	0	100.0	
	PSV2	21	3	85.7	
	PSV3	9	1	88.9	
		36	4	88.9	7.02
Pine, regenerated	R1	15	2	86.7	
	R2	25	2	92.0	
		40	4	90.0	4.15
Pine, cutover	PC1	12	1	91.7	
	PC2	6	0	100.0	
	PC3	21	3	85.7	
		39	4	89.7	3.55
Hardwood, established	H1	35	2	94.3	
	H2	24	0	100.0	
		59	2	96.6	12.47
Hardwood, cutover	HC1	15	2	86.7	
	HC2	8	1	87.5	
	HC3	8	0	100.0	
		31	3	90.3	8.52
Cut/bulldozed	C/B1	77	0	100.0	
	C/B2	84	3	96.4	
		161	3	98.1	1.69
Grass	G1	27	0	100.0	
	G2	21	4	81.0	
	G3	42	2	95.2	
		90	6	93.3	12.63
Bare soil	B1	15	3	80.0	
	B2	12	2	83.3	
	B3	12	0	100.0	
	B4	24	2	91.7	
	B5	30	0	100.0	
		93	7	92.5	4.35
Lakes	L1	30	0	100.0	
	L2	5	0	100.0	
	L3	4	0	100.0	
		39	0	100.0	0.25

Overall Performance, 96.1%
Average Performance by Class, 93.8%

TABLE A-VIII.- STATISTICS OF LARSYS TEST FIELDS

Class Name	No. Points	Misclassified	% Correct	% Area
Pine,	110	9	91.8	
established	112	0	100.0	
	222	9	95.9	45.37
Pine, site	8	0	100.0	
prepared	24	3	87.5	
vegetated	32	3	90.6	7.02
Pine,	8	0	100.0	
regenerated	20	3	85.0	
	28	3	89.3	4.15
Pine, cutover	16	2	87.5	
	16	2	87.5	3.55
Hardwood,	30	1	93.3	
established	30	1	96.7	
	60	2	95.0	12.47
Hardwood,	18	2	88.9	
cutover	18	2	88.9	8.52
Cut/bulldozed	15	1	93.3	
	44	0	100.0	
	59	1	98.3	1.69
Grass	36	2	94.4	
	35	0	100.0	
	35	2	94.3	
	106	4	96.2	12.63
Bare soil	60	0	100.0	
	35	0	100.0	
	95	0	100.0	4.35
Lakes	12	1	91.7	
	12	1	91.7	0.25

Overall Performance, 95.7%
Average Performance by Class, 93.3%

APPENDIX B

GLOSSARY

Accuracy percent – This term has two general uses: to define correctness of delineation along border lines between classes, and the correctness of overall totals of class area. The first is measured by superimposing the classification map on a ground truth map and counting points of agreement versus disagreement between the two maps with a grid. The second is measured by determining the differences in total area of each class between the classification map and the ground-truth map.

Achievement percent – An average between accuracy percent and classification percent.

Band – A group of wavelengths of light which produces one color or convenient grouping of wavelengths, such as near-IR.

Channel – The same as "band" when used in computer work.

CIR film – A three-layer film which images near-infrared as red, red as green, and green as blue.

Classification percent – A measure of the success with which all possible classes are mapped. If only seven of fourteen possible classes are mapped the classification percent is 50 percent.

Clustering – A computer program which groups similar spectral values.

MSS – Multispectral scanner, one of the two basic sensor instruments in the ERTS-1.

Multiband – A study involving more than one band.

Performance percent – This is a measure of the purity or homogeneity of a classification. A count of pixels in a training field reveals the percentage of correct classes versus incorrect. The term is frequently confused with

accuracy, but it is only indirectly related to it. For example, a pine stand may have numerous groups of hardwoods or grassy clearings within it, which are too small to map, and by accuracy standards the area is classed as pine with 100-percent accuracy. However, the non-pine elements appear in the pixel count and may show a low performance percentage.

Pixel — A picture element, the smallest unit recorded by the multispectral scanner.

Signature — A set of spectral, tonal, or spatial characteristics of a classification which serves to identify a feature by remote sensing.

Working circle — A term used by forest managers to describe the territory needed to keep a wood using industry in continuous operation.